Evaluation of statistical cloud parameterizations

Von der Fakultät für Physik und Geowissenschaften

der Universität Leipzig

genehmigte

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt

von **M.Sc. Heiner Matthias Brück**

geboren am 09. Mai 1987 in Hamburg

Gutachter: Prof. Dr. Johannes Quaas
Prof. Dr. A. Pier Siebesma

Tag der Verleihung 24. Oktober 2016

Bibliographic Description:
Brück, Heiner Matthias
Evaluation of statistical cloud parameterizations
Universität Leipzig, Dissertation
183 p., 153 ref., 60 fig., 1 tab.

Abstract:

This work is motivated by the question: how much complexity is appropriate for a cloud parameterization used in general circulation models (GCM). To approach this question, cloud parameterizations across the complexity range are explored using general circulation models and theoretical Monte-Carlo simulations. Their results are compared with high-resolution satellite observations and simulations that resolve the GCM subgrid-scale variability explicitly. A process-orientated evaluation is facilitated by GCM forecast simulations which reproduce the synoptic state.

For this purpose novel methods were develop to a) conceptually relate the underlying saturation deficit probability density function (PDF) with its saturated cloudy part, b) analytically compute the vertical integrated liquid water path (LWP) variability, c) diagnose the relevant PDF-moments from cloud parameterizations, d) derive high-resolution LWP from satellite observations and e) deduce the LWP statistics by aggregating the LWP onto boxes equivalent to the GCM grid size. On this basis, this work shows that it is possible to evaluate the sub-grid scale variability of cloud parameterizations in terms of cloud variables.

Differences among the PDF types increase with complexity, in particular the more advanced cloud parameterizations can make use of their double Gaussian PDF in conditions, where cumulus convection forms a separate mode with respect to the remainder of the grid-box. Therefore it is concluded that the difference between unimodal and bimodal PDFs is more important, than the shape within each mode. However the simulations and their evaluation reveals that the advanced parameterizations do not take full advantage of their abilities and their statistical relationships are broadly similar to less complex PDF shapes, while the results from observations and cloud resolving simulations indicate even more complex distributions. Therefore this work suggests that the use of less complex PDF shapes might yield a better trade-off. With increasing model resolution initial weaknesses of simpler, e.g. unimodal PDFs, will be diminished. While cloud schemes for coarse-resolved models need to parameterize multiple cloud regimes per grid-box, higher spatial resolution of future GCMs will separate them better, so that the unimodal approximation improves.

Extended abstract

This work is motivated by the question: how much complexity is appropriate for a cloud parameterization used in general circulation models (GCM). To approach this question, cloud parameterizations across the complexity range are explored using general circulation models and theoretical Monte-Carlo simulations. Their results are compared with high-resolution satellite observations and simulations that resolve the GCM subgrid-scale variability explicitly.

First, the relationships of the saturation deficit distribution with its cloud properties is explored. It is found that unimodal and symmetric PDFs exhibit a) an unambiguous scaling of liquid water mean (μ_{q_l}) and variability (σ_{q_l}) and b) a PDF-type characteristic value of liquid water dispersion (σ_{q_l}/μ_{q_l}) and skewness for different saturation deficit variabilities, when binned with to cloud fraction. Cloud fraction can therefore be interpreted as a powerful parameter to characterize the PDF's normalized shape. The relationships are extended for the vertically integrated cloud layer and an analytical relationship for liquid water path (LWP) variability is derived on the assumption of maximum vertical cloud condensate overlap and using the column bulk inter-layer condensate correlation.

Second, the range of different complexity cloud parameterizations are explored using GCM simulations for a quasi-hemispheric domain, that covers a all major subtropical stratocumulus regimes and adjacent shallow cumulus regions. The advanced but relatively complex 'unified higher-order turbulent transport- and cloud parameterization' (CLUBB) and the 'eddy-diffusivity dual mass flux scheme' (EDMF-DualM) are explored using GCM forecast simulations. The behavior of the simpler diagnostic schemes, assuming a triangular and single Gaussian PDF, is studied alongside to estimate the benefits of complexity, i.e. physical realism. For this purpose a PDF-moment diagnostic was development and

implemented in the GFDL-AM3 and ICON-GCM. Differences among the PDF types increase with complexity, in particular the more CLUBB and EDMF-DualM make use of their double Gaussian PDF. One can conclude that the difference between unimodal and bimodal PDFs is more important, than the shape within each mode. Using joint histograms of cloud fraction versus cloud properties, two major specialties of CLUBB and EDMF-DualM are revealed. First, high in-cloud σ_{q_l} of shallow cumulus clouds shows a distinct second branch in the phase-space of CF and dispersion, while the majority follows the single Gaussian relationship. Second, EDMF-DualM follows primarily the single Gaussian relationship, but as a result of very small variability for low CF, the scaling is more similar to a uniform PDF.

Third, the role of subgrid-scale temperature variability for cloud parameterizations is explored, which is motivated by the circumstance that most statistical cloud schemes in GCMs only consider subgrid-scale humidity variations, thereby neglecting the influence of temperature fluctuation on saturation humidity. Their joint thermodynamic variation can be studied in the framework of saturation deficit (s). A developed modified version of CLUBB shows that subgrid-scale temperature variability can contribute to saturation deficit variability in the same order of magnitude as subgrid-scale humidity variations. The contribution of subgrid-scale temperature variations are a) directly or b) indirectly via their covariability with humidity variations. Within the marine boundary layer the contribution of the covariability term is larger than the direct temperature variability influence. Neglecting subgrid-scale temperature variability will lead to an underestimation of cloud fraction when CF is less than 0.5 and vice versa to an overestimation for larger CF. However liquid water consistently underestimated. Therefore CF differences that occur between shallow cumulus and stratocumulus regions are further amplified.

Fourth, global high-resolution satellite observations are used to evaluate the subgrid-scale variability produced by GCM cloud parameterizations: This is achieved by a novel approach which is characterized a) by short GCM forecast simulations that capture the synoptic situations and provide the basis for a process-based evaluation and b) by comparing the parameterized vertically integrated subgrid-scale variability directly to corresponding large-scale satellite observations, via aggregating MODIS liquid water path (LWP) observations onto boxes equivalent to the GCM grid size and deriving the LWP statistics in relation to this scale. The conceptual relationships between LWP and cloud cover are quite well captured by the CLUBB and EDMF-DualM, however both schemes feature too extreme cloud cover. Even though CLUBB and EDMF-DualM use a double Gaussian PDF, which could possibly result in broad spectrum of μ_{LWP}-σ_{LWP} combinations, they favor shapes similar to a single Gaussian PDF. In contrast, the LWP-PDF statistics from MODIS show a broader range, that cannot be related to simple PDF-types in a straightforward manner.

Fifth, results from the novel synoptic high-resolution ICON large-eddy model (ICON-

LEM) of the HD(CP)^2 project are presented and analyzed in context to GCM cloud parameterizations. Their ability to explicitly resolve the subgrid-scale variability of GCMs is tested by using the spatial aggregation method already applied to the MODIS observations. ICON-LEM is able to produce realistic cloud cover and integrated condensate variability. The results are slightly larger then corresponding MODIS observations, but more realistic then the benchmark COMSO simulations. The spatial variability is not significantly affected by different regridding resolutions. A direct comparison of the GCM's subgrid-scale cloud variability is complicated by grid-scale differences. Conceptual relationships of LWP-PDF moments are similar to the quasi-hemispheric domain, while ICON-LEM exhibits much larger LWP variability at low cloud cover which can be attributed to smaller, but more condensate loaded shallow cumulus clouds compared to MODIS and COSMO.

The complexity, in terms of assumed PDF shape, of cloud parameterizations is a balance between physical realism, computational effort and the potential to integrate and link ideas within existing GCMs. Approaching the initial question from a conceptual-, GCM modeling-, cloud resolving modeling- and observational point of view, the resolved sub-grid scale variability is more complex than predictable using state of the art cloud parameterizations. Nevertheless this work promotes the use of less complex PDF shapes such as the unimodal, triangular or Gaussian PDF for cloud parameterizations of operational GCMs. This perception is supported by the circumstances that: a) even though higher complexity, i.e. more degrees of freedom, improves some aspects of macroscopic cloud properties their additional benefits might not justify the extra effort; b) higher complexity obfuscates the parameterization's behavior which impedes the attribution of differences deduced from evaluation efforts; c) with increasing model resolution initial weaknesses of simpler, e.g. unimodal PDFs will be diminished. While cloud schemes for coarse-resolved models need to parameterize multiple cloud regimes per grid-box, higher spatial resolution of future GCMs will separate them better, so that the unimodal approximation improves.

List of symbols and abbreviations

Variables

q_t	Total specific humidity	g kg $^{-1}$
θ_l	Liquid water potential temperature	K
RH	Relative humidity	[ ]
q_{sat}	Saturation humidity	g kg $^{-1}$
s	Saturation deficit	g kg $^{-1}$
CF	Cloud fraction	[ ]
q_l	Liquid water content	g kg $^{-1}$
CWP	Cloud water path	kg m $^{-2}$
LWP	Cloud liquid water path	kg m $^{-2}$
SST	Sea surface temperature	K
LTS	Lower tropospheric stability	K
τ	Cloud optical depth	[ ]
CTT	Cloud-top temperature	[K]
CRE	Cloud radiatve effect	W m $^{-2}$

Operators

μ_x or $\bar{x}$	Mean of x	[x]
x'	Fluctuation of x around $\bar{x}$	[x]
σ_x	Standard deviation of x	[x]

γ_x	Skewness of x	[ ]
ν_x	Dispersion of x ($= \sigma_x/\mu_x$)	[ ]
x_g or $x_{grid-box}$	x on grid-box scale	[x]
x_c or $x_{in-cloud}$	x on in-cloud scale	[x]

Abbreviations

CLUBB	Cloud Layers Unified by Binormals
CMIP	Coupled Model Intercomparison Project
COSMO	Consortium for Small-scale Modeling
EDMF-DualM	Eddy-diffusivity dual mass flux scheme
GCM	General circulation model
GFDL	Geophysical Fluid Dynamics Laboratory
GFDL-AM3	Global atmospheric model of the GFDL (third generation)
HD(CP)^2	High Definition Clouds and Precipitation for Advancing Climate Prediction
ICON	ICOsahedral Non-hydrostatic modeling framework
ICON-GCM	Global version of ICON using NWP physics
ICON-LEM	Local area ICON LES model
IPCC	Intergovernmental Panel on Climate Change
ITCZ	Intertropical Convergence Zone
LES	Large-eddy simulation
NWP	Numerical weather prediction
MODIS	Moderate Resolution Imaging Spectroradiometer
PDF	Probability density function
SCM	Single-Column Model

CONTENTS

Chapter 1

Introduction

Clouds are an essential component of the climate system. Their direct influence and their related physical processes strongly shape: a) the earth radiative energy balance via reflection, scattering, absorption and emission of electromagnetic radiation. b) The vertical coupling of the atmosphere via convective re-distribution of energy, moisture and momentum which is driven by the uptake and release of heat. In particular the latent heating associated to water phase changes enables a complex cloud behavior as heat can be stored and transported, which further promotes c) the coupling marine and land regions via atmospheric spatial redistribution that is driven by the interplay of clouds and circulation (Arakawa, 2004; Arakawa and Jung, 2011; Stevens and Bony, 2013; Bony et al., 2015).

The radiative cloud forcing, i.e. the cloud radiative effect (CRE), is largely influenced by the interplay of cloud optical depth (τ) and cloud-top temperature (CTT). A dominating shortwave CRE (mainly driven by τ) will cool the planet and vice versa a dominating longwave CRE (mainly driven by CTT for $\tau > 4$) will warm the planet (Tiedtke, 1989; Hartmann et al., 1992). Since low level clouds have warm CTT they feature a negative net CRE and their ability to cool the planet strongly increase with the optical thickness which can be as large as -200 W/m^2 for thick clouds with $\tau \approx 30$ (Kubar et al., 2007). In order to asses the global radiative impact of clouds, the cloud type specific CRE needs to be weighted with the area covered by clouds, which results in a global mean shortwave effect of -47.3 W/m^2, longwave effect of 26.2 W/m^2 and net effect of -21.1 W/m^2 (Boucher et al., 2013a). In particular regions of persistent cloud regimes such as the stratocumulus, the trade-wind cumulus, the deep convective Intertropical Convergence Zone (ITCZ) and mid-latitude storm tracks show a distinct signature (cloud regimes illustrated in Fig. 1.3).

How clouds might respond to a warming climate is one of the central scientific questions in the earth system science community, in particular whether and how strong a positive cloud feedback will further increase climate sensitivity and potentially amplify global warming (Stephens, 2005; Webb et al., 2006; Bony and Dufresne, 2005; Boucher et al., 2013a). Despite great efforts such as the 'Coupled Model Intercomparison Project' (CMIP) and regular assessments such as the 'Intergovernmental Panel on Climate Change' (IPCC) the spread in climate sensitivity remains however large (Vial et al., 2013; Boucher et al., 2013b). A considerable part of the climate sensitivity uncertainty can be attributed to cloud feedbacks (Zelinka et al., 2012) which can further be traced back to the numerical representation of boundary layer clouds in general circulation models (GCM) (Medeiros et al., 2008; Gettelman et al., 2012; Brient and Bony, 2013; Jakob, 2014; Neggers, 2015a).

Recent advances identified the role of the insufficient representation of heat and moisture mixing in the lower troposphere as a key role. GCMs that re-distribute moisture more efficiently in the vertical, tend to have a higher climate sensitivity since they dry out the lower troposphere, which goes along with reduced low level cloudiness and hence a reduced negative CRE, i.e. a positive cloud feedback (Sherwood et al., 2014; Brient et al., 2015; Tomassini et al., 2015).

A promising idea to improve the understanding of boundary layer clouds is the exploration of the marine trade-wind boundary layer because the large-scale conditions are relatively steady and the interplay of clouds and boundary layer circulation can develop in an unimpeded manner. While boundary cloudiness is well connected to the large-scale atmospheric stability on seasonal time-scales (Klein and Hartmann, 1993; Wood and Bretherton, 2006), relationships on shorter time-scales remain challenging as well as the identification of the key driving mechanisms that control low-level cloudiness (Brueck et al., 2015; Nuijens et al., 2015a;b).

1.1. Subgrid-scale variability

The representation of clouds and their associated effects in GCMs is challenging because of their multi-scale nature. On the one hand microphysical and radiative processes occur on the molecular level, while on the other hand circulation interactions associated with extra-tropical baroclinic systems effect regions up to several thousand kilometers. Therefore no modeling approach can encapsulate the whole cloud phenomena because the numerical discretization of global GCMs or even cloud resolving large-eddy simulations cannot resolve all scales, which in turn needs to be parametrized, i.e. treated in a statistical manner. Cloud related processes on scales smaller than the model grid-box size, i.e. the subgrid-scale, are then described in terms of grid-scale variables. Parameterizations can be interpreted as statistical theories that capture the small-scale processes and describe their interaction with the grid-scale. Their formulation is usually a combination of physical theories and closure assumptions, where the physical realism is balanced against practical modeling constraints. The emergent behavior of cloud-scale interactions creates a certain irony: even though the theoretical equations of small-scale radiative and microphysical processes are well established, their bulk behavior cannot be deduced in a straight-forward manner (Randall et al., 2003).

One particular challenge of this approach is the non-linear nature of the involved processes. Most prominent are the non-linear relationships of cloud radiative and microphysical processes with the grid-scale cloud parameters: e.g. the grid-box mean reflectivity/ precipitation-rate of an inhomogeneous cloud is not the same as the reflectivity/ precipitation-rate that results from the grid-box mean cloud properties (Cahalan et al., 1994; Larson et al., 2001b). Therefore the subgrid-scale variability has to be taken into account explicitly, in order to prevent errors that can be in the same order as the considered effect itself (Pincus and Klein, 2000; Barker and Räisänen, 2004). Even small errors in the grid-scale and subgrid-scale representation of clouds will rapidly amplify and spread trough their dynamical coupling, which makes the parametrization of clouds an essential topic, while their incorporation in GCMs remains challenging (Golaz et al., 2002a).

Over the last decade computing performance increased rapidly, allowing for higher model resolutions. Even though finer grid spacing will resolve parts of the subgrid variability, the effective model resolution (i.e. ability to resolve features) is approximately 7 times the grid spacing, and still the scales of the boundary layer are too fine to be resolved in a foreseeable future so that the core problem remains (Arakawa and Jung, 2011).

1.2. Macroscopic cloud parametrization

The multitude of subgrid-scale cloud processes can be grouped in categories to facilitate their development and implementation in GCMs. However, this pragmatic separation is to some degree artificial since the processes are interconnected. Still, following overarching categories evolved: a) boundary layer transport parametrization which account for the subgrid-scale turbulent- and convective transport of energy, momentum and mass. b) Macroscopic cloud parametrization to predict grid-box cloud fraction (CF) and liquid water content (q_l), and depending on the complexity higher order moments (this category is the topic of this work and will be named 'cloud parameterization'). c) Small-scale cloud radiation and microphysics parameterizations. Typically these cloud related parameterizations are solved sequentially, i.e. the boundary layer transport parametrization supplies essential information to the macroscopic cloud parametrization which in turn provides the input for radiation and microphysics parameterizations.

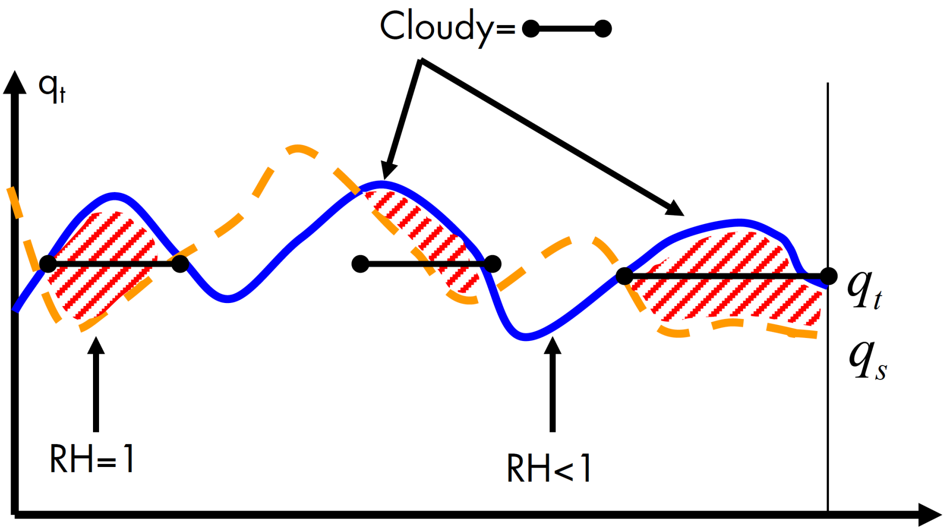


Figure 1.1.: Conceptual figure of subgrid-scale variability and partial grid-box saturation. Adapted from an ECMWF lecture series by Tompkins (2008)

Fractional cloudiness is the essential manifestation of subgrid-scale variability and its prediction is the primary task for a cloud parametrization. The conceptual idea can be understood by considering the spatial variability of the total water (q_t) distribution (vapor+liquid) in relation to the saturation humidity (q_s). Based on the assumption that enough condensation nuclei are available, clouds occur locally whenever the total water value is larger than the corresponding saturation humidity, which is illustrated by the red shaded areas in the one-dimensional model in Fig. 1.1 by Tompkins (2008). Cloud fraction can subsequently be derived by computing the ratio of the cloudy area versus the total grid-box area. The mutual variation of both q_t (blue) and q_s (orange) lines (Fig. 1.1) can be combined by regarding their difference, i.e. the deviation of q_t from saturation (saturation deficit). Considering all saturation deficits (s) along the spatial dimension a distribution function of saturation deficit can be inferred. A systematic approach how to incorporate subgrid-scale variability in cloud parameterizations was initially developed by Sommeria and Deardorff (1977) and Mellor (1977). Based on an assumed probability density function (PDF) of humidity and temperature, cloud fraction (CF) and condensate (q_l) can be derived by integrating over the saturated part of the PDF. Generally all cloud schemes assume directly or in-directly a distribution function of q_t or s and therefore often referred to as statistical cloud parameterizations (even an all-or-nothing cloud scheme can be related to a delta-PDF). Since the pioneering work of Sommeria and Deardorff (1977) and Mellor (1977) various assumed PDF shapes have been proposed, trying to find a balance between complexity (e.g. computational costs and physical insight) and flexibility (e.g. adaptability to different cloud types). A detailed review can be found in Tompkins (2002) and Rosch et al. (2015).

A useful criterion to distinguish cloud parametrization classes is whether they are diagnostic or prognostic: diagnostic cloud parameterizations derive their PDF moments purely from other grid-scale variables and therefore the cloud properties do not exhibit a temporal memory. In contrast the prognostic approach actively carries information of the PDF moments (or cloud properties) in time and space, which requires additional prognostic variables that are advanced by the dynamical core of the model. Thus this approach is more complex as additional equations governing the temporal evolution need to be formulated in terms of grid-scale variables.

1.3. Overview of boundary layer cloud regimes

Throughout this work the behavior of cloud parameterizations will be discussed in context to prototype cloud regime which are therefore introduced here. Clouds are embedded in the atmospheric general circulation. As outlined above the GCM representation of boundary layer clouds is in the focus of climate uncertainty, in particular over subtropical marine regions under large-scale subsiding vertical motions they cover vast domains (Warren et al., 1988). The conceptual relationships are illustrated in Fig. 1.2 which is cited here from van der Dussen et al. (2016): large-scale subsidence is induced by the descending branch of the Hadley circulation and is accompanied by a warming due to the increasing pressure. The associated stabilization of the atmosphere creates an inversion (black line) which marks the upper bound of the marine boundary layer. The lower bound is characterized by gradually increasing sea surface temperature (SST) along the trade-wind trajectory (indicated by the color gradient in Fig. 1.2).

Cumulus convection occurs in the marine boundary layer above the lifting condensation

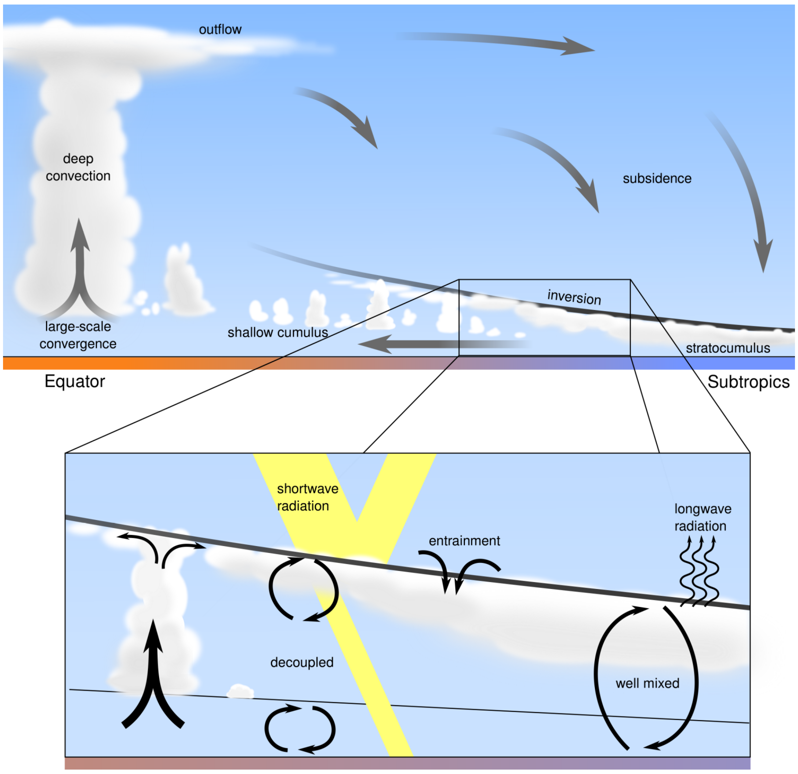


Figure 1.2.: Conceptual illustration of the Hadley circulation and embedded cloud regimes after Arakawa (1975), adapted and modified by Emanuel (1994); Stevens (2005) and cited here from van der Dussen et al. (2016) (graphic published under CC BY 3.0, no modifications here, https://creativecommons.org).

level and four major cloud regimes are prevalent which can be distinguished physically and structurally (Malkus, 1958; Stevens, 2005; Wood, 2012): a) stratocumulus (Fig. 1.3, region 1), b) stratocumulus to shallow cumulus transition (Fig. 1.3, region 2), c) shallow cumulus (Fig. 1.3, region 3), d) deep precipitating cumulus (Fig. 1.3, region 4). In Fig. 1.3 regions of these prototype cloud regimes are indicated on the basis of qualitative visual inspection using satellite imagery. Additionally the quasi-hemispheric domain (145°W to 25°E, 50°S to 60°N) is introduced, which will be referred to in the following chapters.

The stratocumulus regions prevail over the eastern subtropical oceans where subsidence is strongest and SST coldest due to upwelling of cold ocean currents (see Fig. 1.3, region 1). Moisture accumulates between these bounds, forming a vertically well mixed marine boundary layer. CF and q_l are high, leading to strong longwave radiative cooling at the cloud top. The associated thermodynamic destabilization leads to turbulence, resulting in thermodynamic variance production. Warm and dry inversion-air is entrained into the cloud top layer by overshooting turbulent eddies.

The stratocumulus to shallow cumulus transition regions are adjacent to stratocumulus regions (see Fig. 1.3, region 2). This regime is characterized by the formation of shallow cumulus clouds below the stratocumulus, which are forced by the increasing surface fluxes as stratocumulus cloud decks are advected over warmer ocean waters. The shallow cumulus clouds supply the stratocumulus with moisture, thereby acting to maintain the stratocumulus cloud deck by balancing the constant q_l loss at cloud top, due to entrainment of free tropospheric air. Alongside increasing surface fluxes, the large-scale subsidence is weakening which leads to a decoupling of the cloud layer and the well mixed surface layer which promotes more energetic convection that leads to an gradual break-up of the stratocumulus cloud deck (see Fig. 1.2 left part of the lower figure).

Shallow cumulus prevail in the remainder of the subtropical ocean along the trade-wind trajectory (see Fig. 1.3, region 3). The general vertical structure of the lower troposphere can be divided in three layers: a well mixed layer between the surface and cloud base, a conditionally unstable cloud layer and a stable and dry free troposphere above. Air in the cloud layer is a composite of moist air rising out of the mixed layer and dry air originating from the free troposphere. Buoyancy of rising clouds is reduced by lateral entrainment of dry air, which limits the vertical depth of the boundary layer. The temporal evolution is in quasi-equilibrium as dissolving clouds are constantly moistening the cloud layer, leaving a more favorable environment for the next convective towers. Along the trade-wind trajectory the boundary layer thus deepens, which is further stimulated by SST increase and subsidence reduction. The surface driven convective clouds form over a wide size-range, due to the surface to volume ratio dependence of lateral entrainment. Larger, i.e. more voluminous, clouds have a relatively smaller surface area and can reach higher altitudes before they dissolve. The size dependence is captured in the vertical evolution

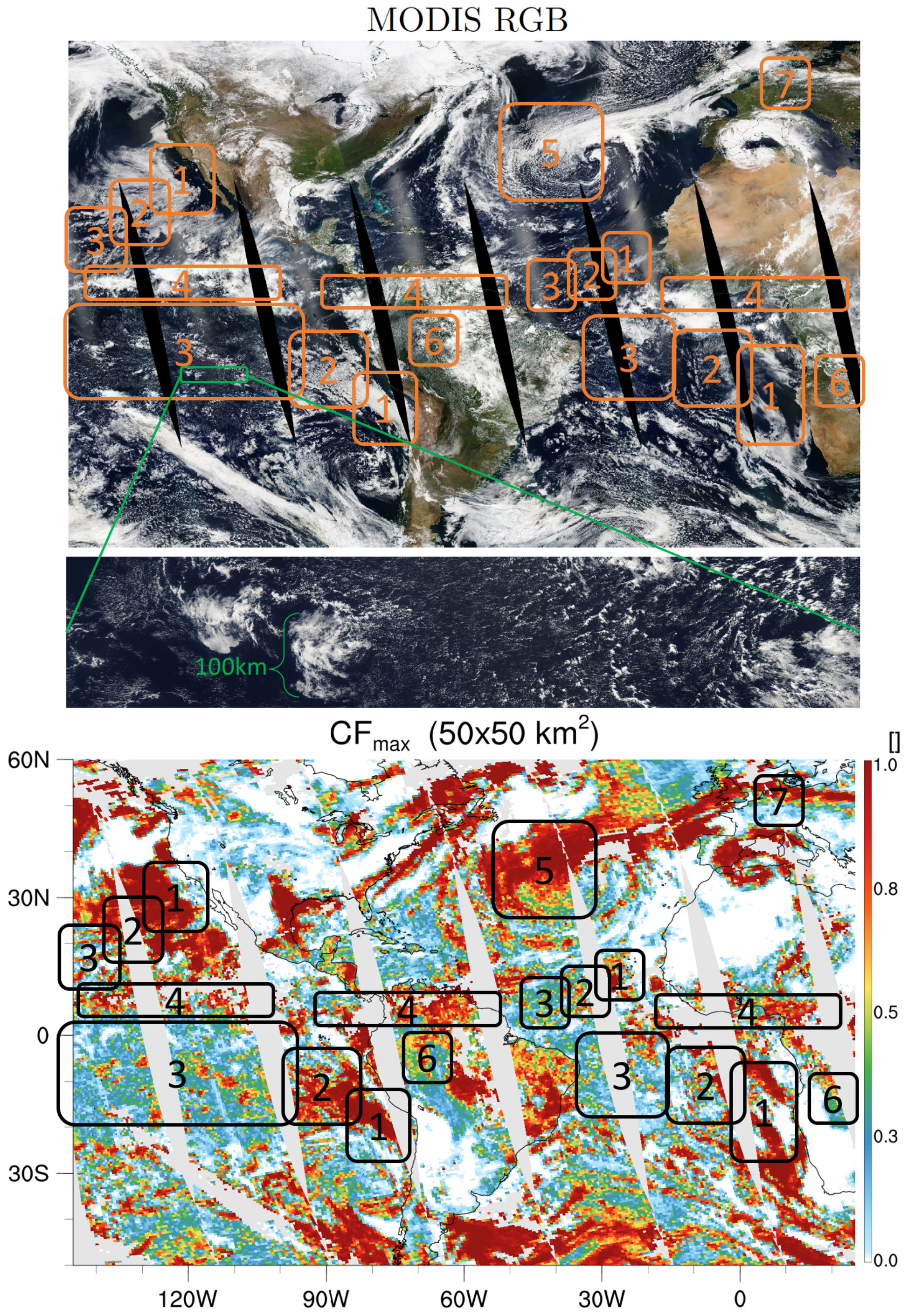


Figure 1.3.: Illustration of regions of prototype cloud regimes based on the inspection of a MODIS visible satellite image and derived cloud cover (not including high level cloudiness, see Chapter 5). The 25 April 2013 is used exemplary. Region 1: stratocumulus; region 2: stratocumulus to shallow cumulus transition; region 3: shallow cumulus; region 4: deep convection; region 5: cold air outbreak; region 6+7: shallow cumulus over land.

of CF and q_l.

The ITCZ is characterized by deep convection which is favored by converging tradewinds and a generally conditional unstable atmosphere (see Fig. 1.3, region 4) which often goes along with the warmest SST. The classical perception that the ITCZ is a continuous band of deep convection, cannot be inferred from the satellite image (Fig. 1.3) and is better characterized by a loosely clustering of deep convective towers that actually do not cover vast regions (Stevens, 2005). Additionally deeper convection occurs in form of isolated organized cloud clusters that are embedded in the shallow cumulus regime, which is illustrated by the magnified satellite image (Fig. 1.3, middle). Notably, less deep and non-precipitating cumulus have the important role of pre-moistening the atmosphere, which in turn supports the development of deeper convection.

Besides the subtropical cloud regimes, the North Atlantic cold air outbreak features boundary layer clouds which are similar to the subtropical stratocumulus-to-shallow cumulus transition regime. An important difference are the much stronger surface fluxes that result from a high near surface temperature gradient when cold air is advected over warm subtropical SST (Fig. 1.3, region 5).

Furthermore shallow cumulus over land will be discussed, which exhibits similar characteristics than marine shallow cumulus but experiences larger sensible heat fluxes that vary in response to the diurnal cycle (Fig. 1.3, region 6,7).

1.4. Scope of this work

This work is motivated by the question: how much complexity is appropriate for a cloud parameterization used in general circulation models (GCM). Over the last decade new advanced cloud parameterizations evolved which make use of double Gaussian PDFs to predict the shape of thermodynamic saturation deficit variability. Therefore one might question how the added complexity relates to simpler PDF shapes in terms of cloud properties and if the benefits justify more complex cloud parameterizations. To approach this question, cloud parameterizations across the complexity range are explored using GCMs and statistical Monte-Carlo simulations. Their results are compared with high-resolution satellite observations and simulations that resolve the GCM subgrid-scale variability explicitly. In particular the following sub-questions are posed in each chapter of this work and their joint proposition is concluded in a last chapter.

- How is the saturation deficit distribution connected to its cloud properties?
- What are the characteristics of complex state-of-the-art cloud parameterizations in GCMs in contrast to simpler ones?
- How important is the subgrid-scale variability of temperature for saturation deficit variability?
- How can global high-resolution satellite observations be used to evaluate subgrid-scale variability produced by GCM cloud parameterizations?
- What can be learned from synoptically realistic cloud resolving models in the context of GCM cloud parameterizations?

Chapter 2

Theoretical relationships of probability density functions

This chapter introduces the conceptual ideas related to cloud parameterizations, which will be further explored using general circulation models and compared in against observations in the following chapters. The conceptual ideas are grouped in tree parts:

First, the saturation deficit concept is introduced which provides a framework to express the joint humidity and temperature subgrid-scale variability. Different probability density functions (PDF) of saturation deficit were used in previous studies (Tompkins, 2002; Rosch et al., 2015) and four of the most commons ones are reviewed and their analytical relationships to the cloud properties are presented and further developed.

Second, to improve the conceptual understanding how the saturation deficit PDF is related its cloud properties, their relationships are explored using Monte-Carlo simulations. Using random generated values that follow these prototype PDFs further allows to validate the analytical relationships established in the first place. The motivation for such an exploration is additionally motivated by the difficulty that simultaneous observations of the total water PDF are very challenging as the water vapor information usually cannot be retrieved in cloudy conditions and vice versa. Therefore a promising idea might be an evaluation approach solely based on the cloudy part of the total water PDF. This idea is motivated by the fact that water vapor and liquid water are tightly connected as they originate from one another.

Third, analytical relationships for the column integrated cloud layer are developed and explored using Monte-Carlo simulations. The motivation is to increase the conceptual understanding how the statistics of the vertically integrated cloud layer are connected to the individual cloud layers. In addition to the second point, the connection the 3D and its 2D statistics would further facilitate the evaluation potential because one could make use of large-scale satellite observations of vertically integrated liquid water, i.e. liquid water path (LWP), which are available over vast domains with horizontal resolutions fine enough to resolve individual clouds.

2.1. Method

Throughout this section the conceptual relationships are explored using analytical equations and Monte-Carlo simulations using random numbers that follow either a uniform, triangular or Gaussian distribution. The random populations are generated using the NCAR Command Language (NCL, 2016). The random library is used internally which follows L'Ecuyer and Côté (1991). The detailed setups of the Monte-Carlo simulations are presented within each section of this chapter.

2.2. Properties of the saturation deficit PDF

Knowing the PDF of humidity and temperature subgrid-scale fluctuations allows to diagnose the cloud properties by integrating over the condensed, i.e. saturated part. This section will explore the four prototype species, i.e. the uniform, triangular, single Gaussian and double Gaussian PDF, that cover a wide range of complexity.

2.2.1. Saturation deficit framework

The thermodynamic state of a grid-box can be characterized by the two conserved variables total specific humidity (q_t) and liquid water potential temperature (θ_l). Their combined effect can be expressed by transforming the thermodynamic state into the saturation deficit framework (Mellor, 1977; Sommeria and Deardorff, 1977; Bougeault, 1981b; Lewellen and Yoh, 1993). Introducing the new variable grid-box mean saturation deficit (s) reduces the number of integration variables to one.

$$s = \frac{\bar{q}_t - \bar{q}_{sl}}{1 + \alpha \frac{L}{c_p}} \tag{2.1}$$

$$\alpha = \left(\frac{\partial q_s}{\partial T}\right)_{T=\bar{T}_l} = \frac{\bar{q}_{sl} L}{R_v \bar{T_l}^2} \tag{2.2}$$

$$\bar{q}_{sl} = \bar{q}_{s(\bar{T}_l)} \tag{2.3}$$

$$T_l = T - \frac{L}{c_p} q_l \tag{2.4}$$

In the equations the constants are: L latent heat of vaporization, c_p heat capacity of air, R_v gas constant of water vapor. In this framework the temperature dependence of saturation humidity (q_s) is linearized around its mean state (Eq. 2.3). A Taylor expansion of saturation deficit allows to express the variability of s (Eq. 2.5), i.e. grid-box mean σ_s, in terms of humidity fluctuations q_t' (term I), temperature fluctuations θ_l' (term II) and their covariability $q_t'\theta_l'$ (term III), using corresponding scaling factors (A,B,C). Following Larson et al. (2002) σ_s is defined as:

$${\sigma_s}^2 = \underbrace{A^2\overline{q_t'^2}}_{\text{Term I}} \underbrace{-2ABC\overline{q_t'\theta_l'}}_{\text{Term II}} \underbrace{+B^2C^2\overline{\theta_l'^2}}_{\text{Term III}} \tag{2.5}$$

$$A = \frac{1}{1 + \alpha \frac{L}{c_p}} = \frac{1}{1 + \beta \bar{q}_{sl}} \tag{2.6}$$

$$B = A\frac{\bar{T}}{\bar{\theta}}\alpha = A\frac{\bar{T}}{\bar{\theta}}\frac{c_p}{L}\beta \bar{q}_{sl} \tag{2.7}$$

$$C = \frac{1 + \beta \bar{q}_t}{1 + \beta \bar{q}_{sl}} \tag{2.8}$$

$$\beta = \frac{L}{c_p}\frac{\alpha}{\bar{q}_{sl}} \tag{2.9}$$

Simpler cloud parameterizations just treat humidity fluctuations explicitly and neglect temperature fluctuations, which potentially cause significant errors. Therefore the special role of temperature fluctuations will be analyzed in Chapter 4. In particular the influence of humidity fluctuations on s-fluctuations can be strengthened by temperature fluctuations, because a negative correlation of $\overline{q_t'\theta_l'}$ produces a positive contribution to σ_s (Eq. 2.5).

2.2.2. Uniform distribution

The uniform PDF is as well denoted as top-hat or box PDF and is unimodal and symmetric. The PDF has equal probability over the interval a to b and is zero otherwise. The shape can be considered as the simplest PDF shape for cloud parametrization besides an all-or-nothing scheme which is represented by delta function. Tompkins showed in an ECMWF lecture series that the uniform PDF is equivalent to critical relative humidity scheme, which was further formalized and explored by Quaas (2012). The critical relative humidity scheme, developed by Sundqvist et al. (1989); Le Treut and Li (1991), is still widely used, often in combination with additional physically motivated 'tweaks', for example in the ECHAM model (Stevens et al., 2013). Following Quaas (2012) and Rosch et al. (2015) the standard deviation can be related to its half-width (Δs), $\Delta s = \sqrt{3}\sigma_s$, and respectively bounding range of (a, b): $a = \bar{s} - \Delta s$ and $b = \bar{s} + \Delta s$. Cloud fraction (CF), grid-box liquid water mean (μ_{q_l} or $\bar{q}_l$) and standard deviation (σ_{q_l}) are defined as:

$$\mathrm{CF} = \begin{cases} 1 & \text{for } a \geq 0 \\ \frac{b}{2\Delta s} & \text{for } a < 0 < b \\ 0 & \text{for } b \leq 0 \end{cases} \tag{2.10}$$

$$\mu_{q_l} = \bar{q}_l = \begin{cases} \bar{s} & \text{for } a \geq 0 \\ \frac{b^2}{4\Delta s} & \text{for } a < 0 < b \\ 0 & \text{for } b \leq 0 \end{cases} \tag{2.11}$$

$$\sigma_{q_l} = \begin{cases} \sigma_s & \text{for } a \geq 0 \\ \sigma_{q_l, a<0<b} & \text{for } a < 0 < b \\ 0 & \text{for } b \leq 0 \end{cases} \tag{2.12}$$

Liquid water path variance on grid-box scale can be derived using a delta function for the clear part $\mathrm{PDF}\left(q_l = 0\right) = \int_{-\infty}^{\infty} \delta_{q_l} \left(1 - \mathrm{CF}\right)\ dq_l$ and $\mathrm{PDF}\left(q_l > 0\right) = \frac{1}{2\Delta s}$:

$$
\begin{aligned}
\sigma^2_{q_l, a<0<b} &= \int_{-\infty}^{\infty} \mathrm{PDF}\left(q_l\right) \left(q_l - \bar{q}_l\right)^2\ dq_l \\
&= \int_{-\infty}^{\infty} \delta_{q_l} \left(1 - \mathrm{CF}\right) \left(0 - \bar{q}_l\right)^2\ dq_l + \int_0^{\bar{s}+\Delta s} \frac{1}{2\Delta s} \left(q_l - \bar{q}_l\right)^2\ dq_l \\
&= \left(1 - \mathrm{CF}\right) \bar{q}_l^2 + \frac{1}{2\Delta s} \left[\frac{1}{3} \left(q_l - \bar{q}_l\right)^3\right]_0^{\bar{s}+\Delta s} \\
&= \left(1 - \mathrm{CF}\right) \bar{q}_l^2 + \frac{1}{6\Delta s} \left[\left(\bar{s} + \Delta s - \bar{q}_l\right)^3 - \left(0 - \bar{s}\right)^3\right] \\
&= \left(1 - \mathrm{CF}\right) \bar{q}_l^2 + \frac{1}{6\Delta s} \left[\left(\bar{s} + \Delta s - \bar{q}_l\right)^3 + \bar{s}^3\right]
\end{aligned}
$$

Relationship of uniform PDF to the critical relative humidity framework

The value 'critical relative humidity' ($\mathrm{RH}_{\mathrm{crit}}$) indicates the minimum grid-box mean relative humidity (RH) for which clouds occur. The lower $\mathrm{RH}_{\mathrm{crit}}$ is, the higher must be the subgrid variability in order to produce sub-regions that are saturated. However RH is bounded to 0, i.e. $\mathrm{RH} = \bar{q}_v / \bar{q}_{sl}$ cannot become negative since the total water $\bar{q}_t$ is positive definite. Hence $\mathrm{RH}_{\mathrm{crit}}$ cannot be negative too.

Following Quaas (2012) CF increases with RH, with a rate determined by $\mathrm{RH}_{\mathrm{crit}}$:

$$
\mathrm{CF} = 1 - \sqrt{\frac{1 - \mathrm{RH}}{1 - \mathrm{RH}_{\mathrm{crit}}}} \tag{2.13}
$$

To reveal the scaling of $\mathrm{RH}_{\mathrm{crit}}$ with RH and CF, Eq. 2.13 is inverted:

$$
\mathrm{RH}_{\mathrm{crit}} = 1 - \frac{1 - \mathrm{RH}}{(1 - \mathrm{CF})^2} \tag{2.14}
$$

Analog, to derive the corresponding uniform PDF variability (σ_s), the half-width of the distribution $\Delta q = \gamma q_s$ with $\gamma = 1 - \mathrm{RH}_{\mathrm{crit}}$ (see Fig. 1 in (Quaas, 2012)) is related to the standard deviation of total water $\Delta q = \sqrt{3}\sigma_{q_t}$ (Rosch et al., 2015). Re-ordering these expressions and using Eq. 2.5 with the assumption of constant temperature across the grid-box:

$$
\sigma_s = A\, \sigma_{q_t} = \frac{A\, q_s}{\sqrt{3}} \left(1 - \mathrm{RH}_{\mathrm{crit}}\right) \tag{2.15}
$$

Note that and the normalization factor A (Eq. 2.6) and the saturation humidity $\bar{q}_s$ depends on the grid-box mean temperature and pressure.

Fig. 2.1 show the scaling of $\mathrm{RH}_{\mathrm{crit}}$ (left using Eq. 2.14) and σ_s (right using Eq. 2.15) in relation to RH and CF using 100 increments for both axis. Temperature and pressure

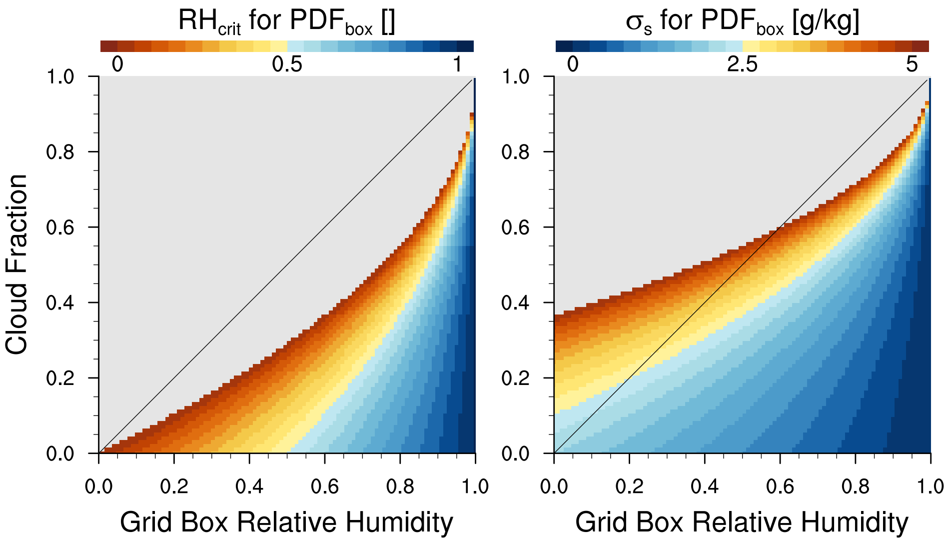


Figure 2.1.: Scaling of $\mathrm{RH_{crit}}$ (left) and σ_s (right) in relation to RH and CF. 100 increments are used for both axis. Grey values indicate missing values, which result from the condition that negative $\mathrm{RH_{crit}}$ is not possible (left) and a subjectively chosen maximum cut-off value for $\sigma_s = 5$ g/kg (right).

are set to 283 K and 900 hPa for illustrative purpose. Other environment states slightly modify the resulting σ_s but do not change the conclusion.

The critical relative humidity framework is only capable to fill a fraction of RH-CF phase-space, i.e. CF is always less than RH. The scaling of CF with RH reaches its maximum for $\mathrm{RH_{crit}} = 0$ and since smaller $\mathrm{RH_{crit}}$ are physically implausible, this approach is bounded. In contrast, this limitation is not present for regular PDF approach where higher CF for a given RH goes along with higher σ_s. Noteworthy in this context is that higher σ_s produces more liquid water which then changes the grid-box mean temperature via latent heat release, which in turn modifies RH and q_s. Generally, using RH to drive cloud parametrization is challenging since it is not directly connected to PDF's position with respect to saturation. Therefore the saturation deficit framework is use subsequently.

The main result of this section is that the critical relative humidity cannot produces high subgrid-scale variability in conditions of low to medium RH, which should be considered when using $\mathrm{RH_{crit}}$ for the evaluation of subgrid-scale variability. In particular regions that exhibit more vigorous moist convection which eventually penetrates into the dry inversion, could exhibit low $\mathrm{RH_{crit}}$ and there is no physical reason why their grid-scale variability should be bounded.

2.2.3. Triangular distribution

Similar to the uniform distribution the triangular distribution is unimodal and symmetric. However the mode is characterized by one value and the tails of the PDF shape are linearly slanted. The triangular PDF has been used occasionally for cloud parametrization purposes e.g. Smith (1990). Following Rosch et al. (2015) the standard deviation can be related to its half-width (Δs), $\Delta s = \sqrt{6}\sigma_s$, and respectively bounding range of (a, b): $a = \bar{s} - \Delta s$ and $b = \bar{s} + \Delta s$.

$$\mathrm{CF} = \begin{cases} 1 & \text{for } a \geq 0 \\ 1 - \frac{a^2}{2\Delta s^2} & \text{for } a < 0 \leq \bar{s} \\ \frac{b^2}{2\Delta s^2} & \text{for } \bar{s} < 0 < b \\ 0 & \text{for } b \leq 0 \end{cases} \tag{2.16}$$

$$\mu_{q_l} = \bar{q}_l = \begin{cases} \bar{s} & \text{for } a \geq 0 \\ \bar{s} - \frac{a^3}{6\Delta s^2} & \text{for } a < 0 \leq \bar{s} \\ \frac{b^3}{6\Delta s^2} & \text{for } \bar{s} < 0 < b \\ 0 & \text{for } b \leq 0 \end{cases} \tag{2.17}$$

The liquid water variability can be derived similarly to Eq. 2.12 but splitting up the integration into the contribution of the right and the left tail. The integration is then analog to Eq. 2.12.

$$\begin{aligned} \sigma^2_{q_l,a<0<b} &= \int_{-\infty}^{\infty} \mathrm{PDF}\left(q_l\right)\left(q_l - \bar{q}_l\right)^2 \, dq_l \\ &= \int_{-\infty}^{\infty} \delta_{q_l}\left(1 - \mathrm{CF}\right)\left(0 - \bar{q}_l\right)^2 \, dq_l \\ &\quad + \int_0^{\bar{s}} \frac{q_l - a}{2\Delta s^2}\left(q_l - \bar{q}_l\right)^2 \, dq_l + \\ &\quad + \int_{\bar{s}}^{b} \frac{b - q_l}{2\Delta s^2}\left(q_l - \bar{q}_l\right)^2 \, dq_l \end{aligned}$$

2.2.4. Gaussian distribution

The Gaussian (or normal) distribution is unimodal and symmetric but continuous, which is different to the uniform and triangular PDF. The slopes are not linearly slanted. Theoretically the probability does not reach zero, which might induce unphysical site-effects for example occasional extreme supersaturations. Sommeria and Deardorff (1977); Mellor (1977) used this PDF to develop the first comprehensive cloud model. Even though the Gaussian PDF is well established in many scientific disciplines, the truncated Gaussian, which results from considering the saturated part, exhibits analytical properties that are mathematically more complex. Still CF and grid-box μ_{q_l} and σ_{q_l} can be analytically derived by making use of the error function (erf).

Assuming a Gaussian distribution of s and integrating over the saturated part, an analytical solution for CF, $\bar{q}_l$ and σ_{q_l} can be derived:

$$\begin{aligned} \mathrm{CF} &= \frac{1}{2}+\frac{1}{2}\mathrm{erf}\left(\frac{\bar{s}}{\sqrt{2}\sigma_s}\right) && (2.18)\\ \mu_{q_l}=\bar{q}_l &= \bar{s}\,\mathrm{CF}+\frac{\sigma_s}{\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\bar{s}}{\sigma_s}\right)^2\right] && (2.19)\\ \sigma_{q_l}^2 &= \mathrm{CF}\sigma_s^2+\bar{s}\;\bar{q}_l\;-\bar{q}_l^2 && (2.20) \end{aligned}$$

The equations follow the derivation of Chen (1991) and the theorems on Gaussian distributions therein. Eq. 2.20 differs from the original Eq. 20 in Chen (1991) and is presented here correctly containing σ_s^2 instead of σ_s.

2.2.5. Double Gaussian distribution

The advanced CLUBB and EDMF-DualM cloud parameterizations (introduced in Chapter 3) use a double Gaussian PDF. The motivation for this more complex shape is to create higher and precisely controllable s-PDF skewness (required in convective regions), which could alternatively be achieved by extending the single Gaussian to a Gamma distribution.

The double Gaussian framework is described by 5 degrees of freedom, two degrees for each PDF and one for their area partitioning. The mean of the envelope PDF that encompasses the double Gaussian PDF can be derived by a simple linear combination using their relative weights. Their joint variance depends on their individual variabilities as well as their respective means. Following Lewellen and Yoh (1993) the first two moments of the joint distribution are:

$$\begin{aligned} 1 &= A_1+A_2 && (2.21)\\ \mu &= A_1\mu_1+A_2\mu_2 && (2.22)\\ \sigma^2+\mu^2 &= A_1(\sigma_1^2+\mu_1^2)+A_2(\sigma_2^2+\mu_2^2) && (2.23) \end{aligned}$$

$$\begin{aligned} \sigma^2 &= A_1(\sigma_1^2+\mu_1^2)+A_2(\sigma_2^2+\mu_2^2)-\mu^2 && (2.24)\\ \sigma^2 &= \underbrace{A_1\sigma_1^2}_{\textit{Term I}}+\underbrace{A_2\sigma_2^2}_{\textit{Term II}}+\underbrace{A_1\mu_1^2+A_2\mu_2^2-\mu^2}_{\textit{Term III}} && (2.25)\\ \textit{Term III} &= A_1\mu_1^2+(1-A_1)\mu_2^2-(A_1\mu_1+(1-A_1)\mu_2)^2 && (2.26)\\ &= A_1\mu_1^2-A_1^2\mu_1^2-2A_1\mu_1\mu_2+2A_1^2\mu_1\mu_2+A_1\mu_2^2-A_1^2\mu_2^2 && (2.27)\\ &= (A_1-A_1^2)(\mu_1-\mu_2)^2 && (2.28) \end{aligned}$$

Using Eq. 2.25 over Eq. 2.24 has the distinct advantage that numerical uncertainties are largely reduced, i.e. the use of Eq. 2.24 might result in a negative variance that cause a problem when deriving the standard deviation (square root), while the transformation yields always positive results.

2.2.6. Relationships of grid-box and in-cloud parameters

To facilitate the discussion and comparison of grid-box scale versus in-cloud scale parameters their relationships will be established. Grid-box scale parameters are typically related to the cloud parametrization itself. Their values are relative to the whole grid-box area, for example cloud fraction indicates the area fraction of the grid-box that is saturated. In contrast, in-cloud scale parameters are defined relative to the cloud area. The discussion of cloud properties in terms of their in-cloud value is often more intuitive. Furthermore the in-cloud properties offer a natural link to other parameterizations, i.e. microphysic and radiation parameterizations, and observations. In other words, in-cloud parameters are independent from cloud fractions which facilitates the comparison among clouds, while grid-scale parameters facilitates the comparison among cloud fields, that contain clear and cloudy sub-domains. Throughout this work in-cloud parameters will be denoted with a subscript x_c or by directly indicating $x_{in-cloud}$ and parameters defined over the cloud free part as x_f, while grid-scale parameters have either no subscript or x_g. The liquid water mean over the cloud area ($\mu_{q_{l,c}}$or $\bar{q}_{l,c}$) is defined as:

$$\bar{q}_{l,c} = \frac{\bar{q}_l}{\mathrm{CF}} \tag{2.29}$$

The liquid water path variance is somewhat more complex to define since the cloud free area is contributing to the grid-box variance too. Making use of Eq. 2.23 and introducing a hypothetical liquid water mean $\bar{q}_{l,f} = 0$ and variance $\sigma^2_{q_{l,f}} = 0$ over the cloud free part, one can define $\sigma_{q_{l,c}}$ as:

$$\begin{aligned}
\bar{q}_l^2 + \sigma^2_{q_l} &= (1 - \mathrm{CF})(\bar{q}^2_{l,f} + \sigma^2_{q_{l,f}}) + \mathrm{CF}(\bar{q}^2_{l,c} + \sigma^2_{q_{l,c}}) && (2.30)\\
\bar{q}_l^2 + \sigma^2_{q_l} &= \mathrm{CF}\ (\bar{q}^2_{l,c} + \sigma^2_{q_{l,c}}) && (2.31)\\
\bar{q}_l^2 + \sigma^2_{q_l} &= \mathrm{CF}\ \bar{q}^2_{l,c} + \mathrm{CF}\sigma^2_{q_{l,c}} && (2.32)\\
\bar{q}_l^2 + \sigma^2_{q_l} &= \frac{\bar{q}_l^2}{\mathrm{CF}} + \mathrm{CF}\sigma^2_{q_{l,c}} && (2.33)\\
\sigma^2_{q_{l,c}} &= \frac{1}{\mathrm{CF}}\left[\bar{q}_l^2 - \frac{\bar{q}_l^2}{\mathrm{CF}} + \sigma^2_{q_l}\right] && (2.34)\\
\sigma^2_{q_{l,c}} &= \frac{1}{\mathrm{CF}^2}\left[(\mathrm{CF} - 1)\bar{q}_l^2 + \mathrm{CF}\sigma^2_{q_l}\right] && (2.35)
\end{aligned}$$

2.2.7. Liquid water skewness

As outlined, the analytical relationships for the first two q_l-PDF moments were established: zero order (CF), first order (μ_{q_l}) and second order (σ_{q_l}). However the analytical solution

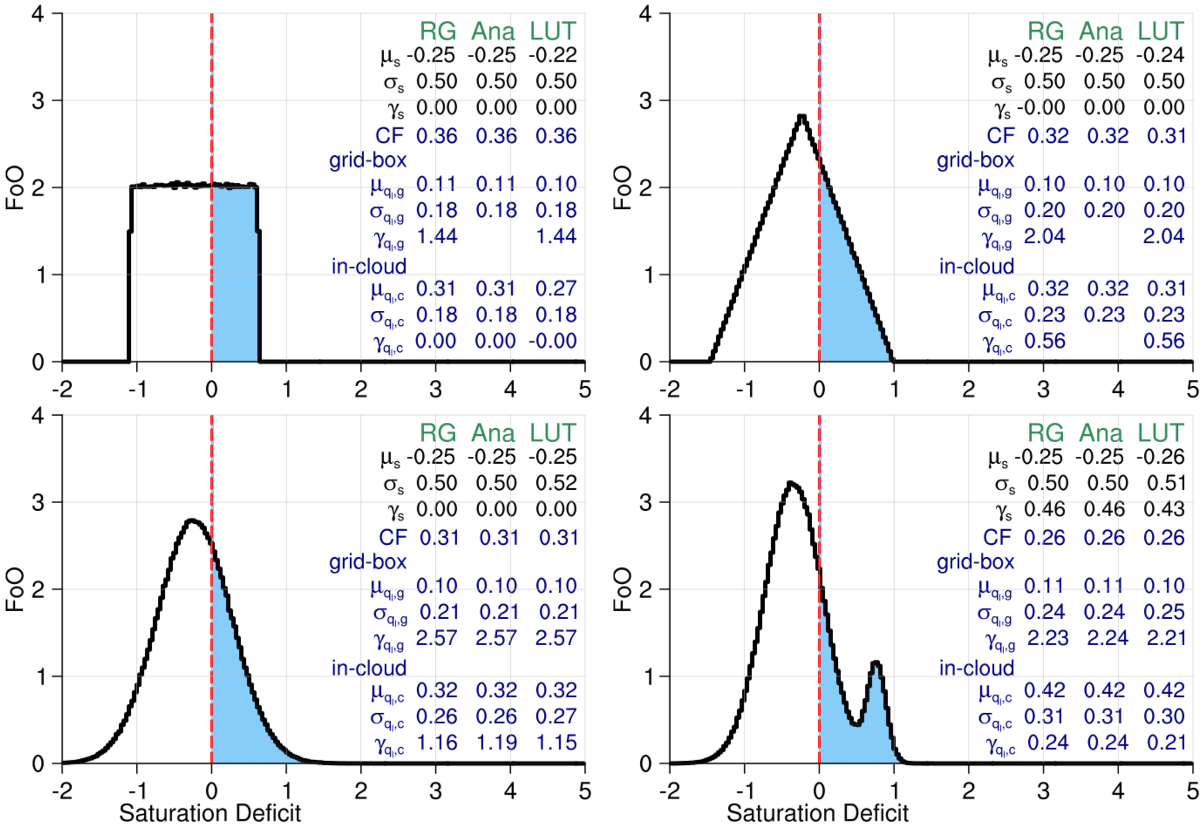


Figure 2.2.: Cloud parameters based on random numbers which are generated assuming a uniform (top-left), triangular (top-right), single Gaussian (bottom-left) and double Gaussian (bottom-right). All distribution functions have the same $\mu_s = -0.25$ and $\sigma_s = 0.5$. The units are arbitrary, but consistent to g/kg. The values inside the sub-figures indicate the moments of the s-PDF, the q_l-PDF relative to the whole grid-box and $q_{l,c}$-PDF relative to the cloud area (in-cloud). The moments are calculated using the random values (RG), the analytical equations (Ana), and are derived from a lookup table (LUT) (see text for further explanation).

for the third order moment, i.e. liquid water skewness (γ_{q_l}), of the Gaussian PDF is theoretically possible, but not straightforward to derive nor to use. Similarly unfeasible is the linear combination of sub-plume γ_{q_l} contribution for the joint double Gaussian PDF. An alternative solution is to use random numbers to create single a Gaussian PDF and subsequently diagnose the γ_{q_l} over the saturated part. For the computation of grid-box scale γ_{q_l}, the cloud free part is set to zero and for in-cloud $\gamma_{q_{l,c}}$, the cloud free part is set to missing value.

2.2.8. Monte-Carlo verification

In order to validate the relationships between the s-PDF and the q_l-PDF moments, a number generator is used to create random values that are distributed according to the four presented PDF types. Deriving the statistical properties purely from random values, offers an independent alternative to the analytical solutions, which is worthwhile as the expressions either contain multiple case switches or are mathematically more advanced (i.e. error function) so that a direct intuitive understanding is not given.

For each PDF type 100000 random saturation deficit values are created. The positive values of the resulting population can be interpreted as the corresponding liquid water values. The fraction of these values is equivalent to the cloud fraction.

Fig. 2.2 shows one possible combination of $\mu_s = -0.25$ and $\sigma_s = 0.5$. The values inside each sub-figure indicate the moments of the respective s-PDF, the q_l-PDF relative to the whole grid-box and $q_{l,c}$-PDF relative to the cloud area (in-cloud). The moments are calculated using the random values (RG), the analytical equations (Ana) and derived from a lookup table (LUT). The LUT is based on cloud properties CF, μ_{q_l} and σ_{q_l}; a concept that will be explored in the next section. The figure is constructed by computing a fine-spaced histogram along the saturation deficit axis. The large number of random samples smoothly retains the underlying distribution functions. Merely the uniform PDF shows slight influence of the randomness.

This hypothetical scenario can be interpreted as a broken cloud-case. Given that all four cases have the same grid-box μ_s and σ_s, it is possible to infer how the different assumed distribution functions effect the CF and q_l-PDF moments:

CF is highest for the uniform PDF and smallest for the double Gaussian, while the average grid-box μ_{q_l} is very similar for all PDFs, hence the in-cloud $\mu_{q_{l,c}}$ is highest for the double Gaussian. Truncating the symmetric triangular and single Gaussian PDFs produce considerable liquid water skewness (γ_{q_l}). Even though the double Gaussian s-PDF is positively skewed, the corresponding q_l-PDF has less skewness than the triangular and single Gaussian PDF. The moderate q_l-skewness results from the second mode of the double Gaussian PDF being completely saturated.

The purpose of Fig. 2.2 is to introduce the concept how the s-PDF is related to the q_l-PDF using one combination of μ_s and σ_s. In the following, these relationships are systematically explored across the physically plausible range of μ_s and σ_s combinations.

2.3. Relationship of cloud and saturation deficit PDF

This section systematically explores how various mutual combinations of μ_s and σ_s influence the resulting q_l-PDF, i.e. the example approach of the previous section (Fig. 2.2) is extended to the physical plausible range. This permutation is done for the uniform, triangular and single Gaussian PDF. These PDFs are unimodal and symmetric so that they only have 2 degrees of freedom which allows to address the various s-PDF shapes in a two dimensional phase-space (in contrast the double Gaussian exhibits 5 degrees of freedom). Fig. 2.3 shows the resulting CF in respect to μ_s for 11 different σ_s, ranging from $\sigma_s = 0$ (blue) to $\sigma_s = 5$ (red). The μ_s ranges from -5 to 5 with a discretization of 500 steps. There is an distinct relation between μ_s, σ_s and CF (except for $\mu_s = 0$ and CF $= 0.5$.). The response of CF to σ_s is different for $\mu_s \pm 0$. For negative μ_s CF increases with σ_s as the right tail of the s-PDF pushes further into the saturated area and vice versa when μ_s is positive. Notably is the relationship between μ_s and CF for different σ_s, which might look quite similar for the three different PDFs at a first glance. However, important differences occur at low CF. On the one hand the single Gaussian is capable to produce low CF for relatively moderate σ_{q_l} compared to the uniform distribution, while on the other hand the single Gaussian needs strong supersaturation to become fully cloudy.

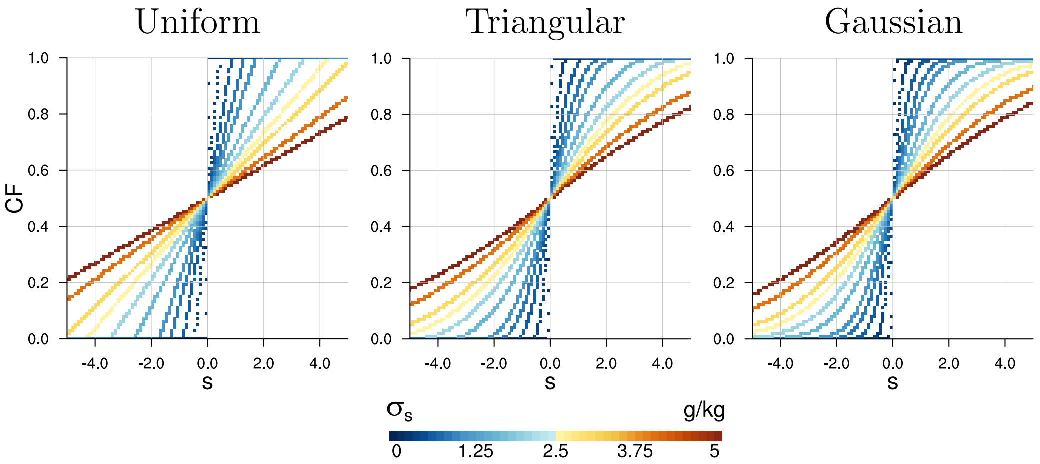


Figure 2.3.: CF in relation to s for a set of different σ_s (colored). A Monte-Carlo approach with N=100000 is used to create a uniform, triangular and single Gaussian s-PDF. μ_s is varied incrementally in 500 steps over the displayed range.

2.3.1. q_l-PDF moments in terms of CF

Even though the underlying s-PDF types are quite simple, their truncated liquid water part exhibits quite complex statistical properties. To study the q_l-PDF moments it is more intuitive to do so in terms of CF. For this purpose the data underlying Fig. 2.3 is rearranged in relation to CF (Fig. 2.4).

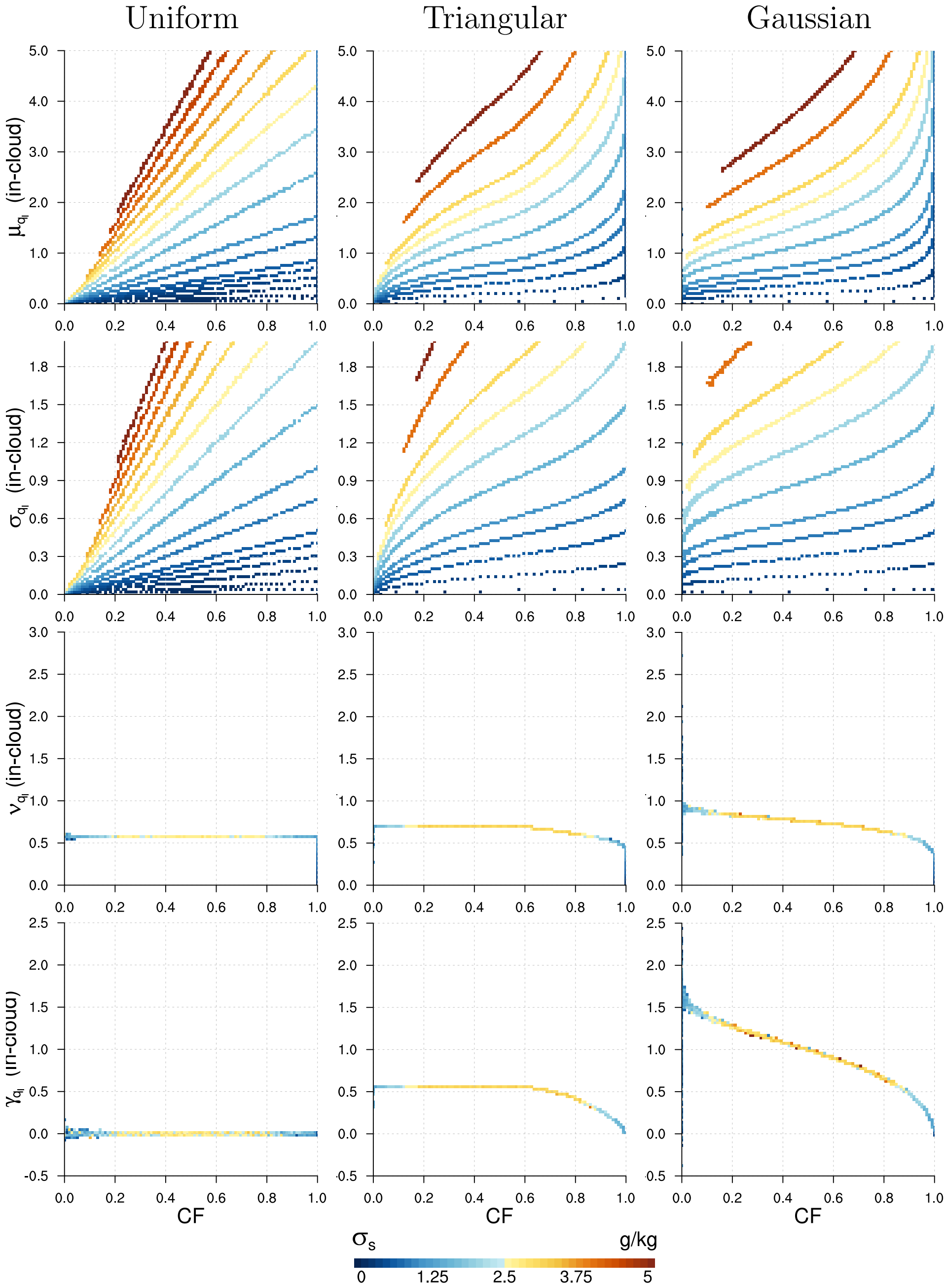


Figure 2.4.: Scaling of liquid water $q_{l,c}$-PDF parameters in relation to cloud fraction (CF) within one layer. The colors indicate the saturation deficit variability (σ_s). The $q_{l,c}$-PDF is computed across a μ_s-range from -5 to 5 with an discretization of 500 steps and a multitude of σ_s. For each μ_s-σ_s combination a Monte-Carlo method (N=100000 values) is used to diagnose the cloud properties. The μ_s-$\sigma_s$phase-space is subsequently rearranged in relation to CF.

$\mu_{q_{l,c}}$ and $\sigma_{q_{l,c}}$ increase monotonically with CF. For the uniform PDF this increase is linear, while the triangular and single Gaussian rise stronger for low and high CF. To facilitate a quantitative description of the q_l-PDF shape the non-dimensional metrics 'liquid water dispersion' ($\nu_{q_{l,c}}$) and skewness ($\gamma_{q_{l,c}}$) are introduced. Dispersion is defined by the ratio of $\sigma_{q_{l,c}}$ divided by $\mu_{q_{l,c}}$ and is a suitable normalized metric of how scattered the distribution is. High dispersion indicates a wide PDF with relatively low mean, while low dispersion is found for narrow PDF with high mean. While $\nu_{q_{l,c}}$ measures the range of variance, skewness of liquid water measures the asymmetry of the q_l-PDF relative to the mode. Positive skewness indicates a tendency of more frequent high water value occurrence than low water ones. Positive $\gamma_{q_{l,c}}$ is generally the case when the q_l-PDF is bounded on the low-end side.

CF corresponds to the truncation point of the underlying s-PDF, regardless of the normalized PDF width. So CF can be interpreted as a normalized measure of shape. In particular σ_{q_l}, $\gamma_{q_{l,c}}$ and their ratio converge to a single value for a given CF. The underlying σ_s has no effect which can be seen from the unique relation of σ_s, $\sigma_{q_{l,c}}$ and $\gamma_{q_{l,c}}$ - CF phase-space (Fig. 2.4). The scaling of $\sigma_{q_{l,c}}$, $\gamma_{q_{l,c}}$ with CF is characteristic for each PDF type. This fact might have important application in order to identify the underlying PDF type. This behavior is found for both q_l-PDFs, i.e. the PDF relative to the whole grid-box where $q_l = 0$ of the clear part are considered as well Fig. A.3 and for the q_l-PDF only over the cloudy area where $q_l > 0$ Fig. 2.4. The scaling of the in-cloud q_l-PDF can be understood more intuitively: The truncated uniform PDF remains a uniform PDF and $\nu_{q_{l,c}}$ is constant across the range of CF. $\nu_{q_{l,c}}$ of a truncated triangular PDF is constant for $0 < \text{CF} < 0.5$ as the shape remains a triangle as long as the mode (i.e. the peak of the PDF) remains unsaturated. However, once the mode of the PDF gets saturated ($\text{CF} > 0.5$) the shape is changing with CF as the portion of the left triangle increases. The single Gaussian is constantly changing with CF. When $\text{CF} = 1$ the q_l-PDF shape can vary freely and so $\nu_{q_{l,c}}$ varies within a wider range. When the s-PDF shape converges to a delta function $\nu_{q_{l,c}}$ becomes 0. In-cloud skewness $\gamma_{q_{l,c}}$ can be understood similarly to $\nu_{q_{l,c}}$, with the exception that $\gamma_{q_{l,c}}$ converges to 0 for $\text{CF} \to 1$ as the non-skewed s-PDF becomes equivalent to q_l-PDF.

Ratio of $\nu_{q_{l,c}}$ and $\gamma_{q_{l,c}}$

The ratio of $\nu_{q_{l,c}}$ and $\gamma_{q_{l,c}}$ (γ/ν_{q_l}) combines the information of the first three moments of the q_l-PDF. Each s-PDF type has a characteristic slope and extent in the $\nu_{q_{l,c}}$-$\gamma_{q_{l,c}}$ phase-space. The Gaussian distribution is theoretically unbounded, which results in occasionally extreme q_l-values. These extreme values can heavily influence the diagnosed q_l-PDF statistics when CF is low (i.e. few q_l values). A similar behavior is introduced from amplified numerical rounding errors when using analytical expressions while CF is low.

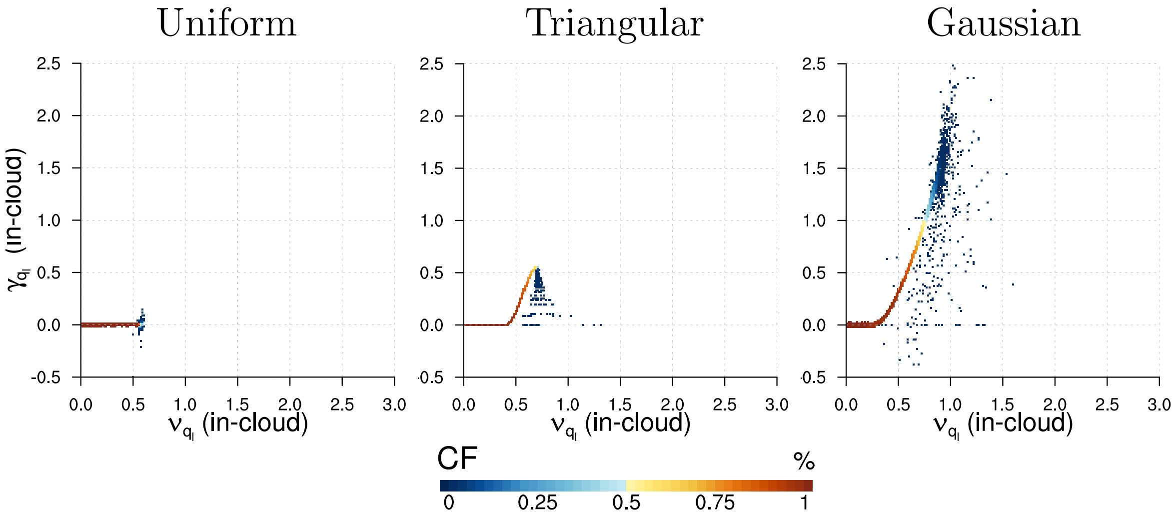


Figure 2.5.: Convergence of the relationship between $\nu_{q_{l,c}}$ and $\gamma_{q_{l,c}}$ for the in-cloud q_l-PDF using data from s-PDF with different σ_s. CF is color-coded. When CF is low, the number of q_l-values produced by the used Monte-Carlo approach is small resulting in incorrect statistics (blue dots which are not along the primary line).

2.3.2. Mutual look-up table

The unique relationships between the s-PDF (characterized by μ_s and σ_s) and the corresponding cloud properties CF, μ_{q_l} and σ_{q_l} facilitates the generation of a look-up table (LUT). The motivation for such an empirical relationship is given by the circumstance that admittedly the derivation of the cloud properties from the s-PDF is possible, but not the other way round, since the underlying equations cannot be inverted (for the Gaussian PDF). Furthermore the unique relationship of γ_{q_l} and ν_{q_l} with CF provides a convenient way to diagnose skewness for unimodal distribution, without using a Monte-Carlo method, which could become computationally expensive. For the generation of the LUT, the concept, underlying the previous section in which μ_s is varied for a set of σ_s, can be extended in a way that both μ_s and σ_s are varied incrementally so that the physically plausible phase-space of CF with the respective cloud variables is filled completely. In principle three permutation of such a 2D-LUT can be realized: CF vs. μ_{q_l} (top row in Fig. A.3), CF vs. σ_{q_l} (second row in Fig. A.3), μ_{q_l} vs. σ_{q_l}. Given that CF is such a powerful parameter in this context, the first two options are explored. The performance of the LUT(CF, μ_{q_l}) and LUT(CF, σ_{q_l}) depends on the gradient of the variable that should be looked up in the respective phase-space. High gradients occur in the CF-μ_s phase-space for low CF, where different σ_s lines are converging. Hence, for the purpose of diagnosing σ_s, the phase-space of CF and σ_{q_l} is more suitable. Furthermore the coverage of this phase-space is higher, which results in lower gradients that lead to higher accuracy of this procedure.

2.4. LWP statistics of the vertically integrated cloud layer

This section explores the vertically integrated cloud statistics that result from multiple cloud layers. Knowing these relationships enables the comparison of 3D fields (GCM, LES) with spatial satellite observations (e.g. MODIS), which will be the analyzed in Chapter 5. The vertically integrated liquid water mean, i.e. the liquid water path (LWP), can be diagnosed straightforwardly by weighting the condensate amounts produced in each layer with the respective layer extend. However the LWP variability cannot be deduced in that manner because the spatial arrangement of condensate within each layer is important. Imagine a checkerboard spatial arrangement in one layer, which is overlayed by a phased-shifted checkerboard pattern. If both layers have equal weights, the LWP variability would be zero, even though the variability within each layer is high. In particular this problem is important for radiation processes since the radiative transfer is highly non-linear related to condensate amounts (Chapter 1).

A general assumption in this context is the independent column approximation (ICA), in which each grid-box column is sub-divided into many smaller columns and the layer cloud condensate is distributed in a manner that the statistical moments are preserved. Radiative fluxes and heating rates are subsequently computed on each sub-column separately and then summed up (Pincus et al., 2006; Barker et al., 2002). Generally there are two main methods to capture this subgrid-scale variability within the framework of ICA: a) using the Monte-Carlo ICA (McICA), which is essentially a random generator to create a statistical set of sub-columns (Räisänen et al., 2004; Weber et al., 2011), or the multivariate sub-column generator e.g. the ‘Subgrid Importance Latin Hypercube Sampler (SILHS)’ recently developed by Larson and Schanen (2013). Or by b), using an analytical vertical integration approach, which is a promising alternative if the considered radiative or microphysical process is relatively simple (Khairoutdinov and Kogan, 2000; Cheng and Xu, 2009; Larson and Griffin, 2013; Griffin and Larson, 2013). Analytical solutions are attractive as they are usually faster and induce less noise compared to the statistical sub-column approach.

In general, vertical integration of subgrid-scale cloud parameters is discussed in the literature in context to radiation or microphysical processes (see citations above). In contrast the scope of this study lies on the macrophysical cloud parameters (like LWP) that result from instantaneous saturation adjustment (Sommeria and Deardorff, 1977; Mellor, 1977). In this sense the use of a full ICA environment might be too complex and obfuscates the vertical superposition, in particular when one assumes maximum vertical overlap of cloud condensate, which simplest case. Motivated by the demand to find a simpler method alongside the goal how to gain deeper understanding how subgrid-scale variability vertically adds up, this section will explore the relationships using idealized cloud layers.

2.4.1. Analytical vertical integral

The vertically integrated liquid water mean over N layers is the weighted linear combination of the individual layer means:

$$\mu = \sum_{i=1}^{N} c_i \mu_i \tag{2.36}$$

The layer weight coefficient c_i is the product of the layer height and average layer density. While the individual layer means are additive, the variances of non-independent random variables is not. Due to the mutual inter-layer correlations of condensate ($\rho_{i,j}$) across all cloudy layers the overall variance of the integrated condensate can be either decreased (negative $\rho_{i,j}$) or increased (positive $\rho_{i,j}$) relative to the simple sum of their variances. The variance of the vertically integrated liquid water is therefore the linear combination of the weighted layer variances plus their mutual covariances:

$$\sigma^2 = \mathrm{Var}\left(\sum_{i=1}^{N} c_i \sigma_i\right) = \sum_{i,j=1}^{N} c_i c_j \,\mathrm{Cov}(\sigma_i, \sigma_j) \tag{2.37}$$

$$= \sum_{i=1}^{N} c_i^2 \,\mathrm{Var}(\sigma_i) + \sum_{i \neq j} c_i c_j \,\mathrm{Cov}(\sigma_i, \sigma_j) \tag{2.38}$$

The diagonal elements ($i = j$) of the covariance matrix are the individual layer variances, while the off-diagonal elements ($i \neq j$) contain all mutual combinations of covariance $\rho_{i,j}\sigma_i\sigma_j$.

Important simplifications can be made for $\rho_{i,j}$ equals 1 and 0. For example (without loss of generality) the case of two layers: setting $c_i = 1$ and using the binomial theorem one can note:

$$\sigma_{1+2}^2 = \mathrm{Var}\left(\sum_{i=1}^{N=2} X_i\right) = \sigma_1^2 + \sigma_2^2 + 2\rho_{1,2}\sigma_1\sigma_2 \tag{2.39}$$

$$\rho_{1,2} = 1 \Rightarrow (\sigma_1 + \sigma_2)^2 \tag{2.40}$$

$$\rho_{1,2} = 0 \Rightarrow \sigma_1^2 + \sigma_2^2 \tag{2.41}$$

Using more than 2 layers will result in multinomials, which can be reduced to binomials using the binomial theorem.

Having a GCM application in mind, one needs to know all mutual inter-layer correlations $\rho_{i,j}$ in order to diagnose the variance of the vertically integrated liquid water from the GCM. One might think that assuming maximum overlap of condensate across the layers does imply a correlation of $\rho = 1$, however when the cloud fraction between two layers is different, their correlation is reduced. Fig. 2.6 shows the saturation deficit PDF and the corresponding q_l-PDF of two exemplary cloud layers. The randomly generated s-values

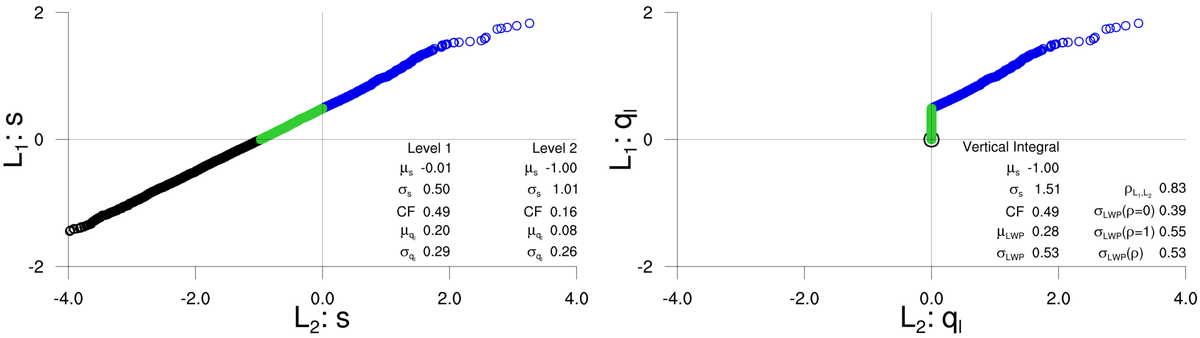


Figure 2.6.: Monte-Carlo realization of a two layer (L_1,L_2) cloud with maximum overlap. 100000 random Gaussian values in each layer. Left: saturation deficit values (s). Right: corresponding liquid water values (q_l). Values which are clear in both layers are in black; cloudy in L_1 and clear in L_2 in green; cloudy in both layers in blue. The statistical moments of the s-PDF and q_l-PDF are indicated in the figure.

follow a single Gaussian distribution. While the s-values line up to a perfect correlation, the truncated q_l-PDF does not. This can be geometrically understood: some of the cloudy points of layer 1 (the more cloudy one) have no corresponding cloudy points in layer 2 (green dots), which effectively prevents an alignment of $\rho = 1$. In this particular example the inter-layer correlation is $\rho_{1,2} = 0.83$. Considering this correlation value in Eq. 2.39 results in the correct cloud integrated variability of $\sigma_{\text{LWP}} = 0.53$ (which is the same as derived from the Monte-Carlo method).

Creating similar realizations to Fig. 2.6, using different s-PDF moments and PDF types reveals that $\rho_{1,2}$ is solely affected by the cloud fraction difference between the two layers, while the shape of the s-PDF in each layer has no influence. This is an important result, however it is not straight-forward to derive an analytical relationship that predicts $\rho_{1,2}$ based on s-PDF shape. Since CF is usually known in a GCM alongside to σ_{q_l}, this concept can be extended to an empirical parametrization of ρ based on a look up table, which will be explored in the following.

2.4.2. Two layer cloud

To gain more insight how the integrated LWP-PDF is affected by the q_l-PDF of individual layers, an idealized two layer cloud is constructed using three combinations of s-PDFs. To create an integrated cloud layer, in which condensate is distributed using the maximum overlap assumption, is as following: First, a Gaussian distributed s-PDF is created using a random number generator. Second, the cloudy part ($s > 0$) is diagnosed and the q_l values are singly sorted (equivalent to the maximum overlap assumption). In the third step, the sorted q_l values of the two layers are summed up and the LWP-PDF moments derived (using an idealized layer weight of 1).

The three idealized s-PDFs have the same $\sigma_{s,2} = 1$ and vary in $\sigma_{s,1} = 1, 0.5, 0.25$. This choice represents the range from homogeneous to inhomogeneous cloud layers. For each combination mean μ_s varies between -5 to 5, i.e. from fully un-saturated to fully saturated. The resulting array in relation to s_1 and s_2 is rearranged in order to display them in terms of CF$_1$ and CF$_2$ (Fig. 2.7). Occasionally there are missing pixels in the CF phase-space (white) which results from a finite resolution of the original $s$phase-space.

Analog to Fig. 2.6 the correlation between the liquid water values ($\rho_{1,2}$) across layer 1 and 2 is derived for all permutations of CF$_1$ and CF$_2$. The major result is a symmetric distribution of $\rho_{1,2}$ in the CF phase-space for all three combinations of σ_s, which highlights that the underlying shape of the s-PDF has no influence. A unique attribution of liquid water values from the two layers occurs when both layers have the same CF which goes along with $\rho_{1,2} = 1$. The contrary case occurs when one layer is fully cloudy and the other one completely clear ($\rho_{1,2} = 0$). There is a general tendency for high $\rho_{1,2}$.

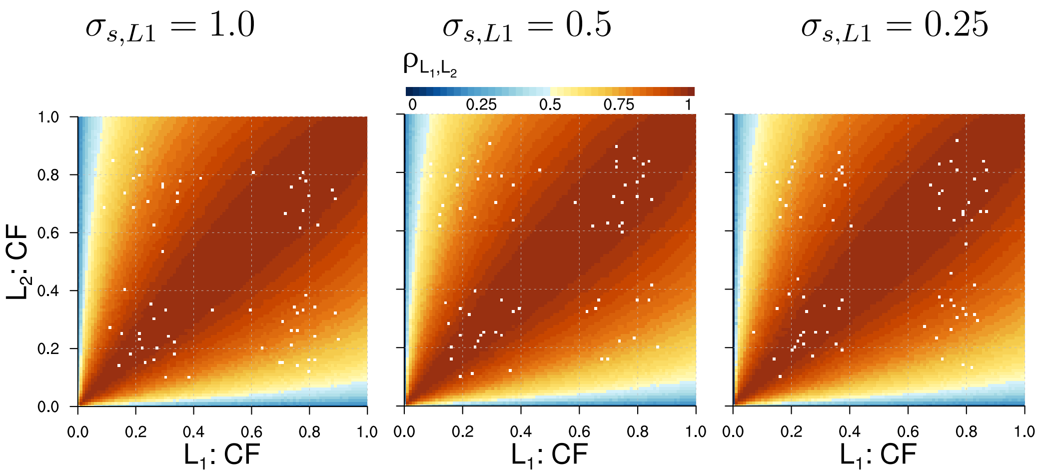


Figure 2.7.: Inter-layer correlation of liquid water values ($\rho_{1,2}$) (color-coded) for three idealized s-PDFs with the same $\sigma_{s,2} = 1$ and different $\sigma_{s,1} = 1, 0.5, 0.25$ (left to right). For each layer μ_s is varied between -5 to 5 and the resulting array in terms of s_1 and s_2 is rearranged to CF$_1$ and CF$_2$.

This invariance of ρ on s-PDF shapes enables a unique diagnostic of ρ in relation to CF$_1$ and CF$_2$, which provides a closure for the integrated liquid water path variance Eq. (2.38) computation. However, the number of cloud layers (N) in a GCM might be larger than 2, which would make this approach tedious, as each $\rho_{i,j}$ would have to be looked up in CF$_i$,CF$_j$ phase-space. Formally, N cloud layers require a (N × N) covariance matrix and one would have to look up half of the off-diagonal elements ($i \neq j$), which are $(N^2 - N)/2$ elements. For example, a GCM might have 20 vertical levels containing liquid clouds, which would result in 190 lookup table calls per column.

A good approximation, computationally more efficient, can be achieved by considering the bulk behavior of the cloud layer in terms of a lookup-table with a ΔCF$_{avg}$ versus CF$_{avg}$ phase-space (Fig. 2.8), in which ΔCF$_{avg}$ is the vertical averaged CF difference among all cloud layer permutations and CF$_{avg}$ the vertical averaged CF. Similarly to (Fig. 2.7),

(Fig. 2.8) the s_1 and s_2 phase-space is rearranged to $\Delta\mathrm{CF}_{avg}$ and CF_{avg}. The vertical average CF can be derived straightforward in the model column, while the column average CF differences requires the use of a covariance matrix too, but each element is the simple difference of two cloud fractions and does not require a lookup-table call. Subsequently a column bulk ρ can be diagnosed form Fig. 2.8 which then can be used in Eq. (2.38) to compute the vertically integrated σ_{LWP}.

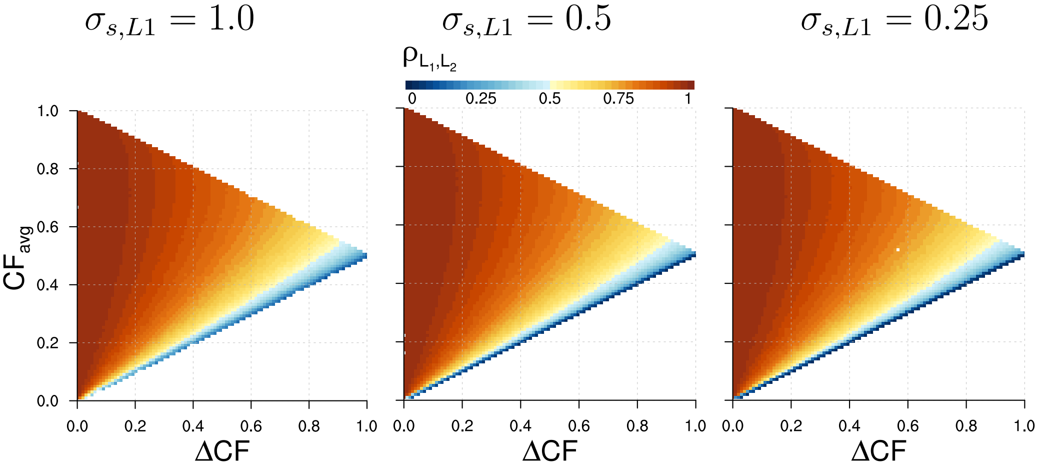


Figure 2.8.: Inter-layer correlation of liquid water values ($\rho_{1,2}$) analog to Fig. 2.7 but rearranged to vertical averaged CF difference among all layers ($\Delta\mathrm{CF}_{avg}$) and vertical averaged CF mean (CF_{avg}). This phase-space can be used as a lookup-table to diagnose the column bulk ρ.

2.4.3. Relationships between cloud cover and LWP-PDF

Following the ideas of the previous section, where the behavior q_l-PDF moments were explored in relation to cloud fraction, this approach is now extended to the scaling of LWP-PDF moments in relation to cloud cover (maximum column cloud fraction). The results from the two cloud layer Monte-Carlo simulations using 3 combinations of σ_s over a wide range of μ_s are used again, but this time rearranged in terms of CF_{max} and respective LWP-PDF moment. Fig. 2.9 shows the relationship of CF_{max} versus μ_{LWP_c}, σ_{LWP_c}, ν_{LWP_c} and γ_{LWP_c} (note this figure is the 2 layer counterpart to the single layer case shown in Fig. 2.4). In contrast to the single layer case the number of free parameters is increasing by two per additional cloud layer (μ_s and σ_s) and eventually by three if the layer weight differs too. The major result is that the relationship of CF_{max} versus ν_{LWP_c} and γ_{LWP_c} is not longer unique. The multitude of combinations produced by two independently varying cloud layers complicates their interpretation, but still there are noteworthy conclusions. First, when $\rho_{1,2} = 0$, one layer is clear and the LWP-PDF moments solely produced by the other layer, which is then equivalent to the respective single layer case. Second, for all combinations of $\sigma_{s,1,2}$ there is the same lower bound for ν_{LWP} and γ_{LWP} at a given CF_{max} for $\rho_{1,2} = 0$. This lower bound is equivalent to the single layer case (in the two layer case

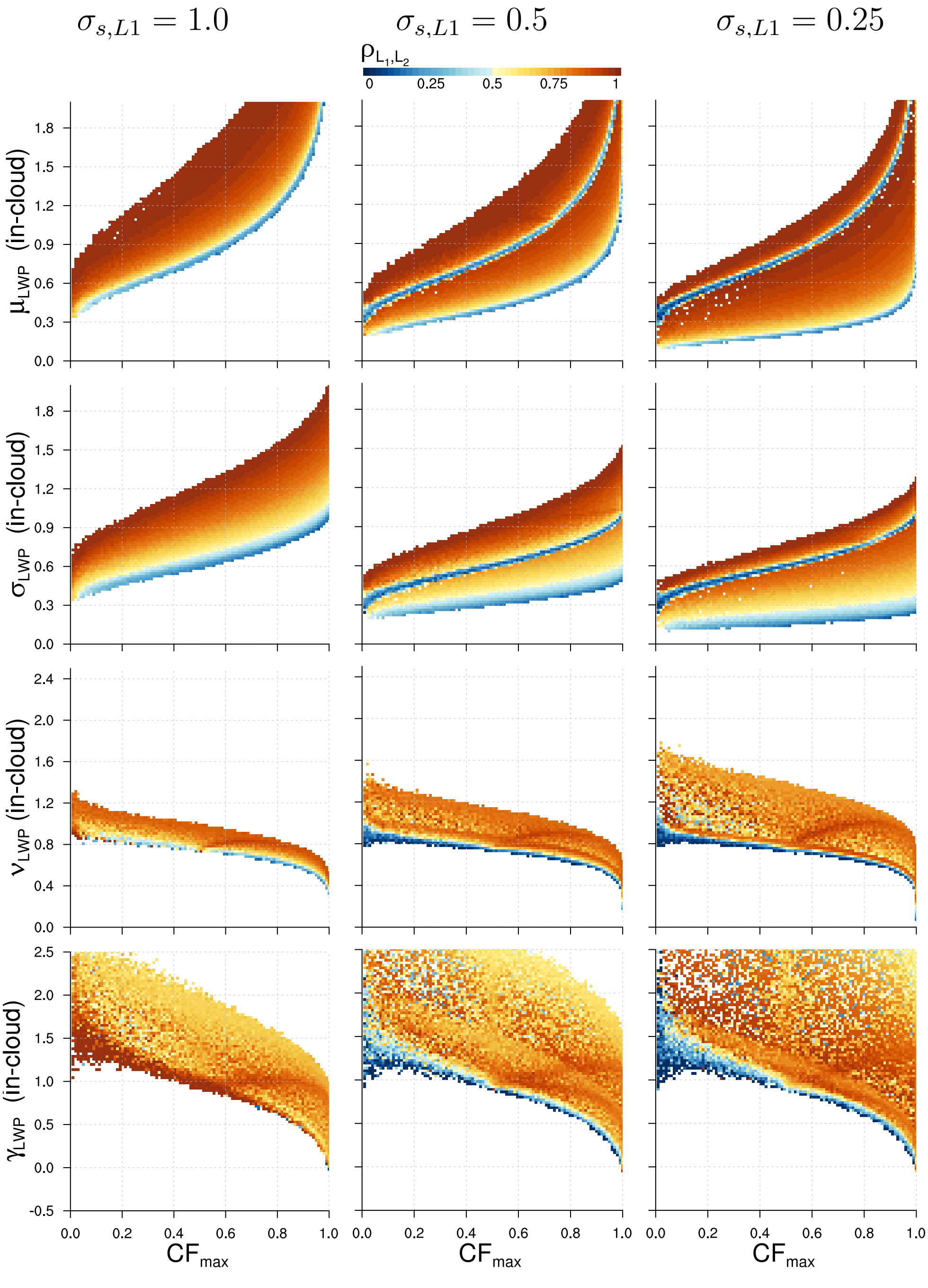


Figure 2.9.: Monte-Carlo simulations: scaling of liquid water LWP-PDF parameters relative to cloud cover (CF_{max}) derived from a two layer cloud. The colors indicate the inter-layer correlation ($\rho_{1,2}$). The LWP-PDF is diagnosed from three s-PDF combinations with the same $\sigma_{s,2} = 1$ and different $\sigma_{s,1} = 1, 0.5, 0.25$ (left to right column). Data rearranged to show the relationships of CF_{max} versus μ_{LWP_c} (row 1), σ_{LWP_c} (row 2), ν_{LWP_c} (row 3), γ_{LWP_c} (row 4) and $\gamma/\nu_{\mathrm{LWP}_c}$ (row 5).

there is somewhat more random noise). Third, the more different $\sigma_{s,1}$ and $\sigma_{s,2}$ are, the larger ν_{LWP}, γ_{LWP}, $\gamma/\nu_{\mathrm{LWP}}$ can get for a given CF_{max}. There is theoretically no upper bound for LWP-PDF moments, when the difference between $\sigma_{s,1}$ and $\sigma_{s,2}$ gets larger.

2.5. Conclusion

Four major prototype probability density distributions (PDF) underlying cloud parameterizations are reviewed in the saturation deficit (s) framework and the analytical solutions for cloud fraction and liquid water mean and variance are presented. Monte-Carlo simulations are used to validate the analytical solutions on grid-box and in-cloud scale and furthermore to explore how a multitude of hypothetical saturation deficit PDFs (s-PDF) translate into cloud properties, i.e. the liquid water PDF (q_l-PDF). Exploring their mutual relationships increases the conceptual understanding and facilitates the interpretation of the cloud parameterization's behavior within GCMs.

Unimodal and symmetric PDFs exhibit an unambiguous scaling of liquid water mean and variability with saturation deficit variability when binned according to cloud fraction. This property is used to create a look-up stable to diagnose the s-PDF moments in terms of the q_l-PDF. A particular special property of unimodal and symmetric PDFs are their unique relationships between cloud fraction and liquid water dispersion (σ_{q_l}/μ_{q_l}) and skewness. For a given cloud fraction, each PDF type produces one characteristic value, independently of the saturation deficit variability. Hence cloud fraction is a powerful parameter to characterize the PDF's normalized shape. These results support the idea, that an evaluation of the cloud parameterization's subgrid-scale variability can be based on the cloudy part. Such an approach would have intriguing capabilities, because the cloudy part of the spatial subgrid-scale variability is better observed than the water vapor part (du to higher spatial resolutions).

In this line of motivation, analytical relationships for the column integrated cloud layer are developed and explored using Monte-Carlo simulations for an idealized two layer cloud model of different saturation deficit PDF combinations. Connecting the 3D cloud layers with their integrated statistics further facilitates the evaluation potential because one could make use of vertically integrated liquid water (liquid water path) of scanning satellite observations, which are available over vast domains with horizontal resolutions fine enough to resolve individual clouds. The subgrid-scale liquid water variability of multiple vertically staggered grid-boxes is not additive (in contrast to the mean) because the spatial overlapping of condensate needs to be considered. Traditionally sub-column generators are use to calculate radiative and microphysical processes that depend on the vertically integrated liquid water variability. However this approach adds additional complexity and induces noise from the finite statistical sample. In this work the possibility of an analytical solution for the integrated liquid water variability was explored. The use of a covariance matrix, that contains all inter-layer covariances, provides a solution. Assuming maximum vertical overlap of cloud condensate, a methods to parameterize the inter-layer condensate correlations was established. Using the Monte-Carlo simulations, it was found that the inter-layer condensate correlations only depends on the cloud fraction difference between the layers. The results can be extended to multiple cloud layers and the use of a column bulk inter-layer condensate correlations promotes the application of this approach in GCMs.

Chapter 3

Cloud related subgrid-scale variability in GCMs

This section explores the simulated subgrid-scale relationships between the saturation deficit (s) probability density function (PDF) and the associated cloud properties. The behavior of recently developed sophisticated cloud parameterizations is explored using free-running GCM simulations that can produce complex interactions with neighboring cells and with the circulation they are embedded in. This approach challenges idealized single-column modeling studies which are usually used alongside large-eddy simulations to develop cloud parameterizations. However it remains unclear if the conceptual ideas can be reproduced in a more complex environment.

Both cloud schemes tested here, utilize a double Gaussian PDF which can theoretically produce complex interactions between the underlying saturation deficit PDF and the resulting cloud properties (Sec. Theory). In relation to simple uniform distributions one might ask what the benefits of this added complexity are? Following this question the simple diagnostic PDFs that have been previously explored using Monte-Carlo simulations, are now diagnostically coupled to both schemes.

Throughout this section grid-box and in-cloud scale cloud parameters will be explored simultaneously, which is motivated a) by a close link of grid-box properties to the cloud parametrization and b) by the more intuitive understanding of in-cloud properties. Furthermore in-cloud properties are potentially more attractive for the coupling of subgrid information to other parameterizations that address microphysical and radiative processes.

The outline of this section is: first, the CLUBB and EDMF-DualM cloud parameterizations are introduced alongside their hosting models GFDL-AM3 and ICON-GCM. The setup of the model simulations and further post-processing steps are described. Second, the results from CLUBB using the GFDL simulation are analyzed and third, followed by the exploration of EDMF-DualM simulation in the same manner. Additionally, in both latter sections the results obtained from simpler cloud schemes are studied and compared. In the fourth section the conceptual behaviors of CLUBB and EDMF-DualM are contrasted with each other.

3.1. Methods

In contrast to the previous chapter where idealized Monte-Carlo simulations were studied, here realistic GCM simulations are performed. The approach is split up the two different cloud parameterizations, i.e. CLUBB and EDMF-DualM, and their respectively host models, i.e. GFDL-AM3 and ICON-NWP.

3.1.1. GFDL-AM3 and CLUBB

This study uses the third generation of the atmospheric model (AM3) (Donner et al., 2011) which is embedded in the third generation of the coupled general circulation model (CM3) developed by the Geophysical Fluid Dynamics Laboratory (GFDL), i.e. GFDL-CM3. The other components of the coupled model environment are not necessary here, since interactions on climatological timescales are not the scope of this study.

GFDL-AM3 is based on a hydrostatic dynamical core using a finite-volume discretization. To avoid polar singularities and better grid uniformity, a cube-sphere grid is used which consists of 6 surfaces, each exhibiting $193 \times 193 = 37249$ columns. The non-orthogonal gnomonic projection then creates a horizontal grid resolution between 40 and 58 km. The are 48 vertical layers between the surface and the model top at 1 hPa. Below 700 hPa the grid stretching results in 12 layers, which is less than in ICON, but still enough to capture essential boundary layer features such as the boundary layer inversion (Donner et al., 2011).

The unified higher-order turbulent transport- and cloud parameterization ‘Cloud Layers Unified by Binormals’ (CLUBB) (Golaz et al., 2002a; Larson and Golaz, 2005; Larson et al., 2012) was successfully implemented in the atmospheric component of GFDL-AM3 (Donner et al., 2011) and serves as an optional unified parameterization for future climate simulations in GFDL-AM3 CLUBB (Guo et al., 2014) and GFLD-AM3 CLUBB+ (Guo et al., 2015). Our simulations follow the GFLD-AM3 CLUBB+ setup, in which CLUBB replaces and therefore unifies the following parameterizations which are otherwise the standard GFLD-AM3 setup: deep convection (Donner et al., 2001), shallow convection (Zhao et al., 2009), cloud macrophysics (Tiedtke, 1993) and planetary boundary layer scheme (Lock et al., 2000).

CLUBB

In CLUBB the unification of the turbulence and convection parameterizations are achieved by predicting a joint PDF of humidity, temperature and vertical velocity for each grid-box and advancing the joint PDF prognostically in time. A double Gaussian for each variable is used to generate this joint PDF. Assuming a double Gaussian PDF is based on Larson et al. (2002) and motivated by the PDF's ability to generate symmetric and skewed distributions (positively or negatively) as well as bimodal ones. 15 PDF parameters define the shape of the three double Gaussian PDFs in w, q_t and θ_l. To determine this framework 10 prognostic moments and 5 additional assumptions are used, which are described in detail in Larson et al. (2002); Golaz et al. (2002a) and Larson and Golaz (2005). The additional closure assumptions mainly address the subgrid-scale distribution of within-plume correlations. The prognostic moments are the first- and second order moments of vertical motion (w), q_t and θ_l, i.e. mean $\overline{w}$, $\overline{q_t}$, $\overline{\theta_l}$ and variances: $\overline{w'^2}$, $\overline{q_t'^2}$, $\overline{\theta_l'^2}$ and their covariance $\overline{q_t'\theta_l'}$, and turbulent fluxes: $\overline{w'\theta_l'}$ $\overline{w'q_t'}$. Additionally the third-order moment $\overline{w'^3}$ is predicted. The area partitioning, i.e. the relative weight of each single Gaussian PDF, is done using the first 3 moments of the vertical velocity PDF (Larson et al., 2002; Larson and Golaz, 2005). This increased complexity makes the scheme computationally more expensive than traditional parameterizations. The idea of a central joint PDF approach separates CLUBB from conventional eddy diffusivity and convective mass flux schemes and offers an intrinsic advantage: higher-order turbulent moments, buoyancy terms, cloud fraction (CF) and q_l can be diagnosed consistently from the same PDF. Therefore the prediction of these terms can be avoided, which limits the number of closure assumptions and makes it easier to couple more complex microphysics and radiation schemes to the cloud properties.

PDF diagnostic

The double Gaussian PDF moments are diagnosed in terms of saturation deficit. However, directly including them into the standard model output in not trivial. First, CLUBB uses a different vertical grid convention than the host model GFDL-AM3 and second, CLUBB uses a nested time-step which is required for a stable solution of the higher-order vertical velocity moments (Golaz et al., 2002a). For this purpose the GFDL-AM3 model source codes was extend. This was achieved in consistent manner, i.e. the mutual relationships between the s-PDF and the q_l-PDF remain the same.

GFDL-AM3 CLUBB+ setup

The free-running GFDL-AM3-CLUBB+ forecast simulations were initialized from a multi-year baseline simulation, which is nudged to ‘National Centers for Environmental Prediction’ (NCEP) reanalysis data (Saha et al., 2010) using the 3D horizontal wind components, water vapor, surface pressure and SST with a relaxation timescale of 6 hours. Direct initialization of forecast simulations from NCEP fields produced very similar results. However this approach proved to be not as robust as the latter because some simulations occasionally terminated when horizontal gradients became too high.

3.1.2. ICON GCM and EDMF-DualM

The ICON (ICOsahedral Non-hydrostatic) modeling framework (Zängl, 2013) was jointly developed by the German Weather Service (Deutscher Wetterdienst, DWD) and the Max Planck Institute for Meteorology (MPI-M) for weather prediction and climate applications. The most important new modeling capabilities are: the non-hydrostatic dynamical core that enables high-resolution cloud resolving simulations, the two-way nesting capabilities which allow regional refinement and the option to choose between a global and limited-area mode. Based upon this new functionality several new models were established from this unified modeling framework: a) a general global circulation model which succeeds the ECHAM GCM within MPI-M (Stevens et al., 2013) b) a global numerical weather prediction model which succeeds the global GME (Majewski et al., 2002) and regional ‘Consortium for Small-scale Modeling’ (COSMO) model within DWD, and c) a large-eddy simulation (LES) model (Dipankar et al., 2015).

The dynamical core of ICON (Wan et al., 2013; Zängl et al., 2015) solves the compressible non-hydrostatic equations which are necessary for a realistic treatment of convection in high-resolution simulation. In order to facilitate high spatial resolutions and to make use of rapidly increasing computing performance of multi-core architectures, a particular emphasis was attributed to the model’s computational performance, i.e. the scalability and efficiency of simulations using $\mathcal{O}(10^4 - 10^6)$ cores. Alongside this development, a new unstructured triangular model-grid was introduced, which is based on the refinement of a spherical icosahedron (R1B0). In this study the ICON R2B05 grid will be used, in particular each triangle of the spherical icosahedron is initially root divided into 2 sections (R2) and further partitioned into 5 bisection steps (B5). This splitting results in a total number of $20 \cdot 2^2 \cdot 4^5 = 81920$ grid-columns. The effective grid resolution is 78.9 km which is defined as the square root of the average cell area (Zängl et al., 2015). The use of equal size triangles removes the long-standing polar singularity problem of regular latitude-longitude grids. Compared to spectral transformation models, which allow analytical solutions of spatial derivatives using Fourier and Legendre transformation, the scalability of global grid cell-to-cell communication is improved (Staniforth and Thuburn, 2012).

The parameterized physics can be grouped according to the respective model scopes. The global climate version follows the ECHAM physics (Stevens et al., 2013), the NWP version is strongly inspired by the COSMO (Doms et al., 2011) and 'Integrated Forecasting System' (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF) physics and the LES version uses the newly developed 3D Smagorinsky turbulence scheme (Dipankar et al., 2015; Heinze et al., 2016). In this chapter simulations using the NWP physics are performed and in Chapter 6 the results from simulations using the LES physics are analyzed. The major NWP physics parameterizations are: the mass flux Tiedtke/Bechtold convection scheme (Bechtold et al., 2008), the EDMF-DualM boundary turbulence and transport scheme (Neggers et al., 2009), the cloud microphysics scheme (Doms et al., 2011), the Rapid Radiative Transfer Model (RRTM) radiation scheme (Mlawer et al., 1997; Barker et al., 2003), the non-orographic gravity wave drag (Orr et al., 2010), the subgrid-scale orographic drag (Lott and Miller, 1997), and an updated version of the COSMO land scheme TERRA (Heise et al., 2006).

EDMF-DualM

The most significant deviation from the setup used here, compared to the default combination of parameterizations, is the choice of the eddy-diffusivity dual mass flux scheme (EDMF-DualM), instead of the prognostic Raschendorfer turbulence scheme based on turbulent kinetic energy (Doms et al., 2011). The EDMF-DualM replaces the parameterization of the shallow convection, while above the default mass flux Tiedtke/Bechtold convection scheme is still active. This additional turbulence and transport scheme has been implemented into the ICON model (Martin Köhler, personal communication) as a complementary option for research purposes in the framework of HD(CP)^2 and the Hans Ertel Centre for Weather Research (HErZ) which promotes ongoing development (Sakradzija et al., 2015).

The EDMF framework evolved over the last years with important contributions by Neggers et al. (2004); Soares et al. (2004); Köhler (2005), comprehensively summarized by Siebesma et al. (2007). The key idea of the EDMF is to unify the boundary layer turbulence and convection as they are deeply connected to each other. Two probability density functions are used to decompose the eddy-diffusive and the convective transport. While the diffusive transport captures the small scale turbulent mixing process, the convective transport represents the more organized and non-local vertical rearrangement. Within a grid-box, convective updrafts occur over a wide range of intensities which resemble the cloud size spectrum. Their joint effect is captured by the bulk mass flux assumption (Tiedtke, 1989). Neggers et al. (2009) extended this framework by splitting up the bulk mass flux into a dual mass flux (DualM). This separation into a dry and a moist updrafts can be considered as the simplest realization of a multi-plume mass flux scheme (a more complex multi-updraft version was recently developed by Neggers (2015b)). The dry mass flux represents the boundary layer plumes that do not reach the lifting condensation level (LCL), while the moist updrafts (the more energetic plumes that reach the LCL and use the extra buoyancy from the latent heating) penetrate into the cumulus cloud layer. Essentially the ideas underlying the EDMF-DualM scheme capture our conceptual understanding of the major physical feedback mechanism of boundary layer clouds that were outlined in Chapter 1. The area fraction of dry and moist updrafts is flexible, which allows a smooth varying between boundary layer cloud types and thus facilitates the gradual transition from stratocumulus to shallow cumulus induced by gradually increasing surface forcing. In EDMF-DualM this area partitioning is parameterized in terms of moist convective inhibition above the mixed layer top, i.e. cloud base (Neggers et al., 2009).

Neggers (2009) further extended the EDMF-DualM framework into a statistical cloud scheme which was integrated deductively by the double moment PDF for the diffusive and convective transport. Each component is transformed into a conserved variable space (q_t vs. θ_l) and prognostically associated with a variance that reflects the subgrid-scale fluctuations. This approach results in a unique position and orientation of each PDF which consistently captures the very nature of the turbulent diffusive and the convective transport process. Cloud fraction and liquid water can then be diagnosed from each PDF by integrating over their respective condensed part using the relationships established by Sommeria and Deardorff (1977); Mellor (1977), and the double Gaussian extension by Lewellen and Yoh (1993) (Chapter 2).

The variance of the moist updraft PDF is inferred from a rising plume model using two test-parcels. Their thermodynamical spread along the vertical trajectory is linked to the variance. The variance of the dry updraft is implicitly included into the diffusive PDF variance. Both are diagnosed from the residual that results from the grid-box mean variance minus the moist updraft PDF variance using the double Gaussian relationships (Lewellen and Yoh, 1993) or Eq. 2.23. Grid-box mean variance is parameterized using a prognostic variance budget which includes flux gradient production, vertical transport and dissipation, while the horizontal components of the variance flux are neglected (Neggers, 2009).

For the purpose of this work, the double Gaussian PDF related moments are diagnosed from the EDMF-DualM routine after they are updated within a time-step. The cloud closure routine of the EDMF-DualM implementation uses a vector notation to address the variability in of the conserved variables q_t and θ_l. Here the double Gaussian PDF is diagnosed in terms of saturation deficit, which are consistent with the associated cloud properties. Furthermore the output routine of the hosting ICON model was extended to include these PDF moments simultaneously to the standard model output.

ICON GCM setup

The ICON GCM simulations are initialized using IFS analysis data at 00 UTC. Using the ICON modeling framework this model capability is an attractive side-effect of shared infrastructure. The stand-alone utility tool 'iconremap' (Prill, 2014) was used to perform area-weighted horizontal remapping from the regular IFS grid to the ICON R2B05 grid. The vertical interpolation is handled by ICON itself in the initialization phase at model startup. The initialization data contains essentially the whole atmospheric and surface state at model start (Prill, 2014). The subsequent spin-up time is relativity short, which is facilitated by the similar physics parameterization used in the IFS and ICON model. After 12 hours the globally averaged boundary layer structure is in a new equilibrium. For this work the second day (24h-48h) is then used.

3.1.3. General setup

For both models the vertically integrated CLUBB and EDMF-DualM cloud properties are computed using the analytical approach developed in Chapter 2 based on the column bulk inter-layer correlation of mutual CF differences. The model output analysis focuses on the quasi-hemispheric domain (145°W to 25°E, 50°S to 60°N) since this particular subset contains all boundary layer clouds regimes of interest including all major subtropical stratocumulus regions (see Fig. 1.3 of Chapter 1). Usually the model output is regridded to a latitude-longitude grid. However, since cloud fraction is one of the key properties in this study, this regridding approach might introduce undesired consequences. For example regridding methods that use spatial interpolation could create artificial cloud fraction values that were possibly never produced by the cloud parameterization. Or even more unwanted, the consistent non-linear relationships between the saturation deficit PDF and the associate cloud properties are perturbed. Therefore I developed data processing methods that are capable of handling the unstructured triangular ICON- and the cube-sphere GFDL grid, which is technical realized using of the Earth System Modeling Framework (ESMF) incorporated in the NCAR Command Language (NCL, 2016).

3.2. GFDL CLUBB

For each grid-box the envelope PDF that encompasses the double Gaussian is constructed (μ_s, σ_s and γ_s) and explored alongside the underlying single Gaussian PDFs, which provides physical insight how the different cloud regimes are parametrized. By convention PDF$_1$ represents the more convective PDF that usually has a higher mean s.

The mutual relationships of the s-PDF and q_l-PDF as well as the vertically integrated LWP-PDF statistics produced by the CLUBB cloud parameterization are explored. In order to preserve cloud regimes of a given synoptic situation, daily averages based on 3 hourly output of the second day of a forecast simulation with the GFDL-AM3 model are

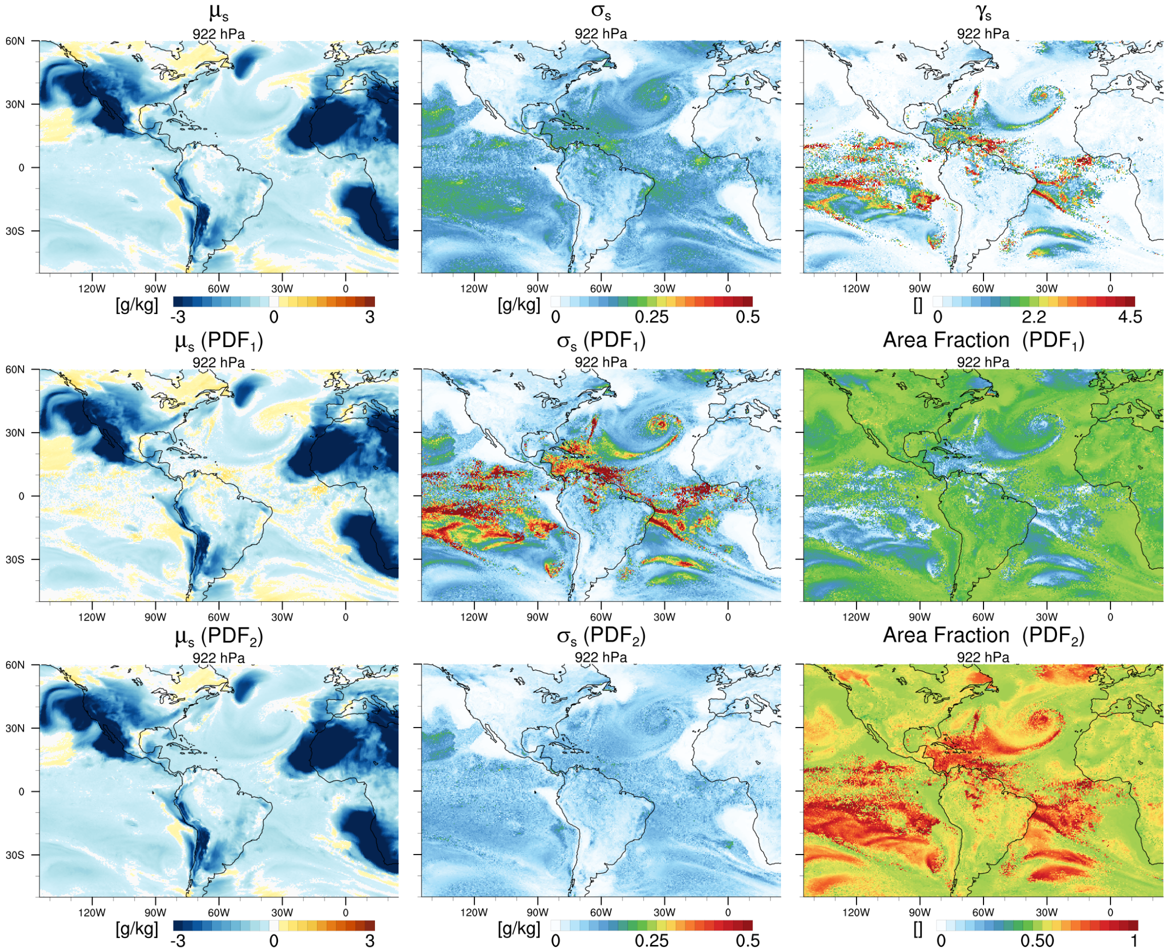


Figure 3.1.: Daily average based on 3 hourly output of the second day of a forecast simulation for 25 April 2013 with the GFDL-AM3 model using the CLUBB cloud parameterization showing the characteristics of the double Gaussian saturation deficit (s) PDF. Top row: grid-box mean μ_s (left), standard deviation σ_s (middle) and skewness γ_s (right). γ_s is almost entirely positive definite. Middle row: contribution from PDF$_1$ and area fraction, which is the relative weight of this PDF component (right). Bottom row: contribution from PDF$_2$. The model level that exhibits the highest stratocumulus cloud fraction is shown. The corresponding average marine subtropical pressure height is 922 hPa.

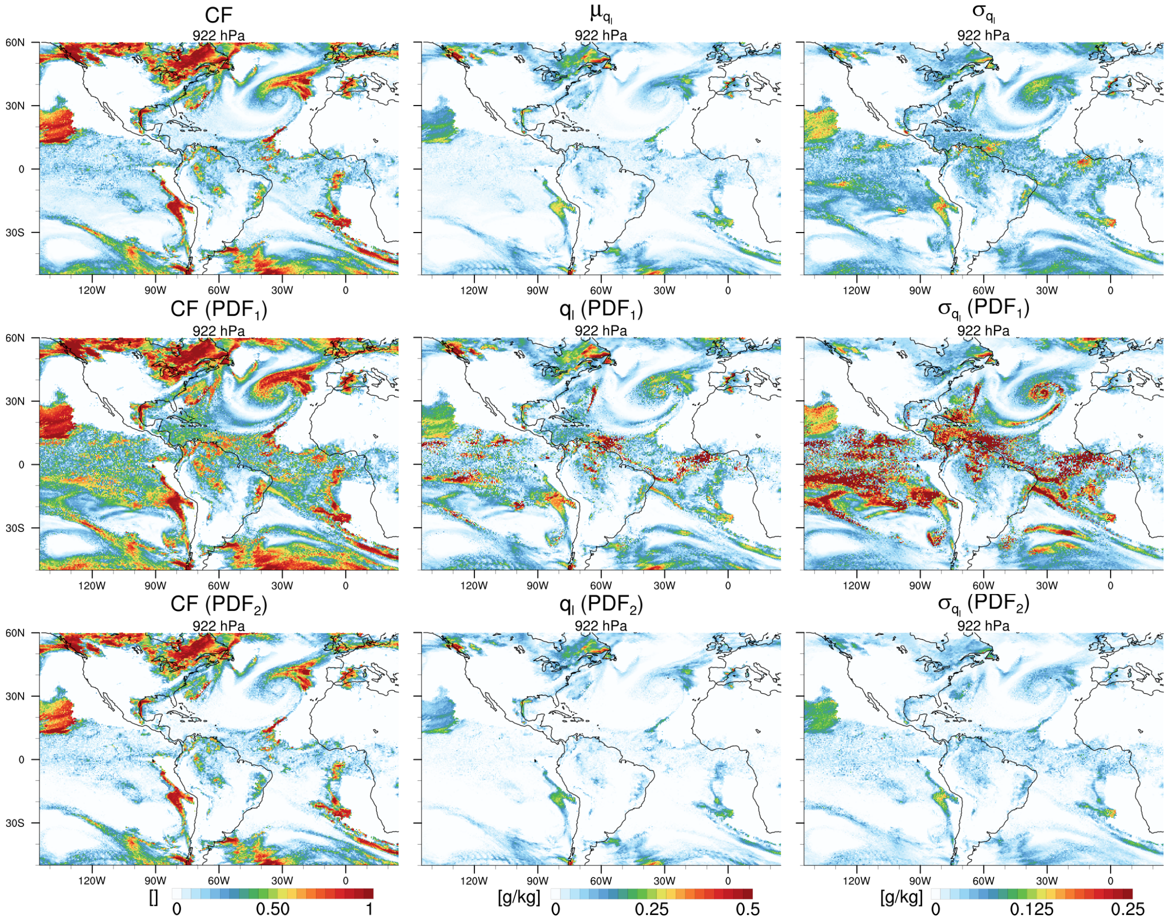


Figure 3.2.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: characteristics of the double Gaussian cloud PDF. Top row: grid-box cloud fraction CF (left), liquid water mean μ_{q_l} (middle) and standard deviation σ_{q_l} (right). Middle row: contribution from PDF$_1$. Bottom row: contribution from PDF$_2$. Note that CF $= \mathrm{A}_1\mathrm{CF}_1 + \mathrm{A}_2\mathrm{CF}_2$ and equivalently for σ_{q_l}.

used. The spatial distribution of the s-PDF (Fig. 3.1), the q_l-PDF (Fig. 3.2) and the grid-scale and in-cloud scale q_l statistics (Fig. 3.3) is shown for the model level that exhibits the highest stratocumulus cloud fraction, which corresponds to an average marine subtropical pressure height of 922 hPa. In conjunction the vertically integrated statistics of cloud properties (Fig. 3.4) and liquid water path (LWP) PDF statistics (Fig. 3.5) are presented. The vertical integral is derived for the lower troposphere (surface up to 650 hPa) which captures the majority of the boundary layer clouds. Subsequently, their relationships are discussed in context to their cloud regime specific behavior which were introduced and geographically referenced in Fig. 1.3 in Chapter 1.

Marine stratocumulus

Marine stratocumulus prevails over the eastern subtropical oceans and is generally represented in CLUBB by one highly cloudy model layer (Fig. 3.4). Both single Gaussian

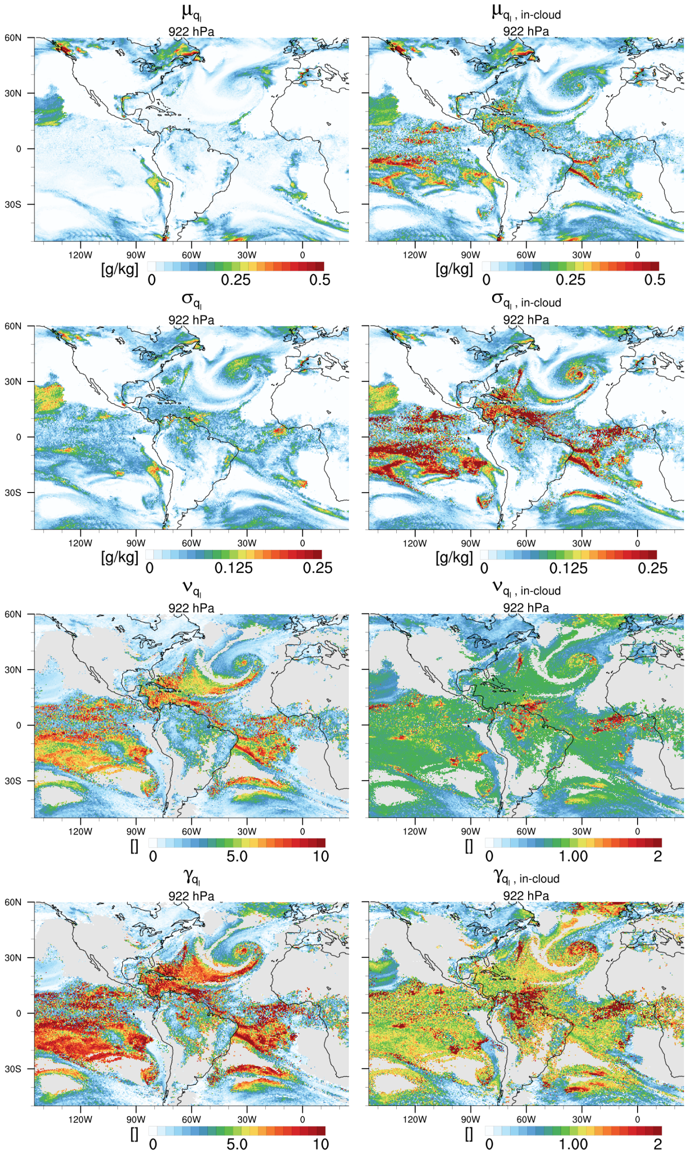


Figure 3.3.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: grid-box mean liquid water PDF statistics on grid-box scale (left) and on in-cloud scale (right). Liquid water: mean μ_{ql} (row 1), standard deviation σ_{ql} (row 2), dispersion $\nu_{ql} = \sigma_{ql}/\mu_{ql}$ (row 3), skewness σ_{ql} (row 4). Grey color indicates cloud free regions, i.e. missing values.

PDFs are relatively similar in terms of μ_s, σ_s, which results in moderately positive grid-box mean s (Fig. 3.1). Horizontal gradients in s are large and negative values indicate regions above the boundary layer inversion within the free troposphere. PDF$_1$ is slightly more saturated but smaller in coverage. The associated s skewness is close to zero and infrequently negative (note the color range of γ_{ql} in Fig. 3.1 is truncated at 0). Negative skewness would occurs when non-cloudy downdrafts merge into the stratocumulus cloud top layer (driven by radiative and evaporative cooling), however this generally not simulated because mean s of PDF$_1$ > PDF$_2$. Both PDFs exhibit moderate and relatively

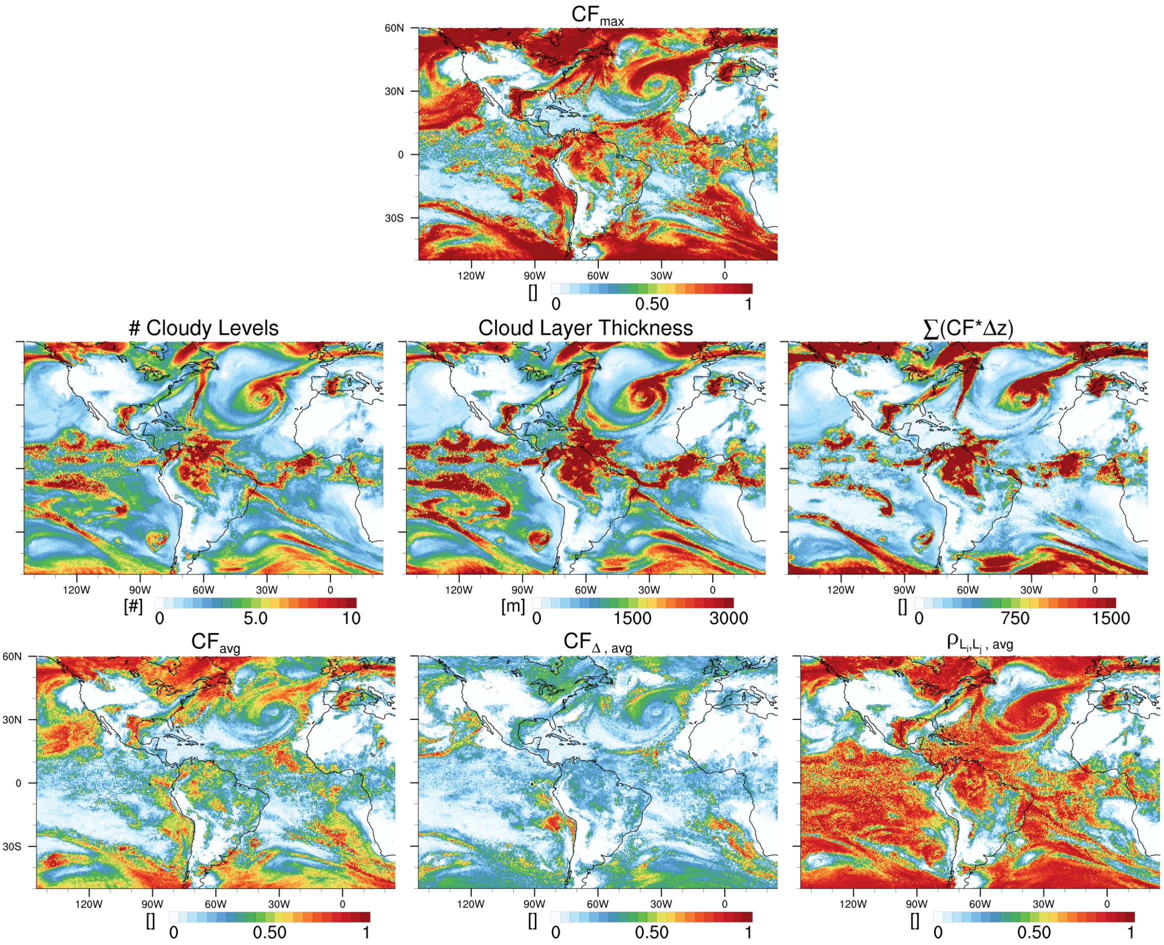


Figure 3.4.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: vertically integrated cloud characteristics. Top row: Maximum cloud fraction (CF) within a column. Middle row: number of cloud layers within a column (left), cloud layer thickness (middle), sum of the layer product of CF times layer thickness (right). Bottom row: average column CF (left), average inter-layer CF difference (middle), bulk inter-layer correlation coefficient (right).

spatially homogeneous σ_s compared to other cloud regimes and PDF$_2$ of the stratocumulus regime is slightly higher, which supports its relative importance for stratocumulus. Horizontal homogeneity of the cloud decks decreases when more than one model layer is cloudy. In that case cloud fraction (CF) between both layers is quite different (Fig. 3.4, bottom-middle). The stratocumulus cloud deck is only partly visible in any individual layer of the 3D data (Fig. 3.2) as the cloud layer rises stepwise along the trade-wind trajectory, while the column projection shows the full extent (Fig. 3.4). Within the layer of maximum cloud fraction, both PDFs exhibit high CF, the mean liquid water (μ_{q_l}) is relatively high, as well as the liquid water variability on grid-box scale (σ_{q_l}). Considering the in-cloud liquid water statistics $\sigma_{q_{l,c}}$, $\gamma_{q_{l,c}}$, $\nu_{q_{l,c}}$ one can note that they are smaller than for other cloud regimes (Fig. 3.3). The LWP statistics of vertically integrated cloud layer are very similar since only few layers contribute (Fig. 3.5).

Marine stratocumulus transition

The transition is characterized by a change from marine closed cell stratocumulus (high cloud cover) to open cellular convection (medium cloud cover) to shallow cumulus convection (low cloud). The transition is initiated by shallow cumuli developing below the stratocumulus decks due to increasing surface fluxes as sea surface temperature increases along the trade-wind trajectory (de Roode and Duynkerke, 1996). The shallow cumuli eventually penetrate into the stratocumulus which over time breaks up the cloud deck (Chapter 1). The representation of this process is a long-standing problem in GCMs and the model results presented here show this shortcoming as well. The south-east Pacific stratocumulus transition is very abrupt, while in the north-east Pacific stratocumulus deck remains closed until rapidly transitioning to deep convection in the vicinity of the Intertropical Convergence Zone (ITCZ). Still, the stratocumulus transitions in the North and South Atlantic are more typical (Fig. 3.4). The transition is accomplished within the double Gaussian PDF framework by μ_s of PDF$_1$ remaining weakly positive, while its area fraction decreases with increasing convective activity, which is reflected by the gradual increase in σ_s of PDF$_1$. μ_s of PDF$_2$ is moderately negative with small μ_s, but produces small cloud amounts with little water variability.

Marine shallow cumulus

Marine shallow cumulus is characterized by low cloud cover and high positive skewness of s. The long positive PDF-tail produces consequently high positive q_l skewness (γ_{q_l}). μ_s of PDF$_1$ is slightly negative, but still larger than PDF$_2$, the corresponding σ_s of PDF$_1$ is high compared to the stratocumulus regime. In that context the area fraction of PDF$_1$ is inversely related to σ_s, which is reflected in the resulting cloud properties where σ_{q_l} is strongly connected to σ_s. Shallow clouds extend over multiple model levels. Their condensate amount is low relative to the grid-box scale, but high for the in-cloud scale. The in-cloud dispersion $\nu_{q_{l,c}}$ (ratio of $\sigma_{q_{l,c}}$ divided by $\mu_{q_{l,c}}$, introduced in Chapter 2) is quite constant and liquid water skewness is moderate.

One particular property of CLUBB stands out: the vertically integrated in-cloud μ_{LWP_c} and σ_{LWP_c} are exceptionally high over vast shallow cumulus regions (Fig. 3.5), along with very high grid-scale γ_{q_l} (Fig. 3.3). This behavior can be explained by the excessive σ_s of PDF$_1$. Even though the area fraction of these circumstances is very small (around 1%) the contribution of this component dominates the envelope PDF and the resulting cloud properties.

Deep convection

The CLUBB scheme handles deep convection along with shallow convection, which is quite a novel feature compared to other boundary layer turbulent transport schemes (Guo et al., 2015). Deep convective regions can be identified by a vertical extent of cloud parameters over many vertical layers (Fig. 3.4). Here only the integral over the lower troposphere is shown. Deep convection occurs in form of frontal structures associated to mid-latitude storms, along the ITCZ in form of large cloud clusters and occasionally more isolated, embedded in the shallow the cumulus regime. LWP mean and variance are high. Interesting to note is the characteristic signature of more isolated convective clusters in terms of in-cloud liquid water dispersion and skewness, which are much higher than boundary layer clouds and synoptic systems.

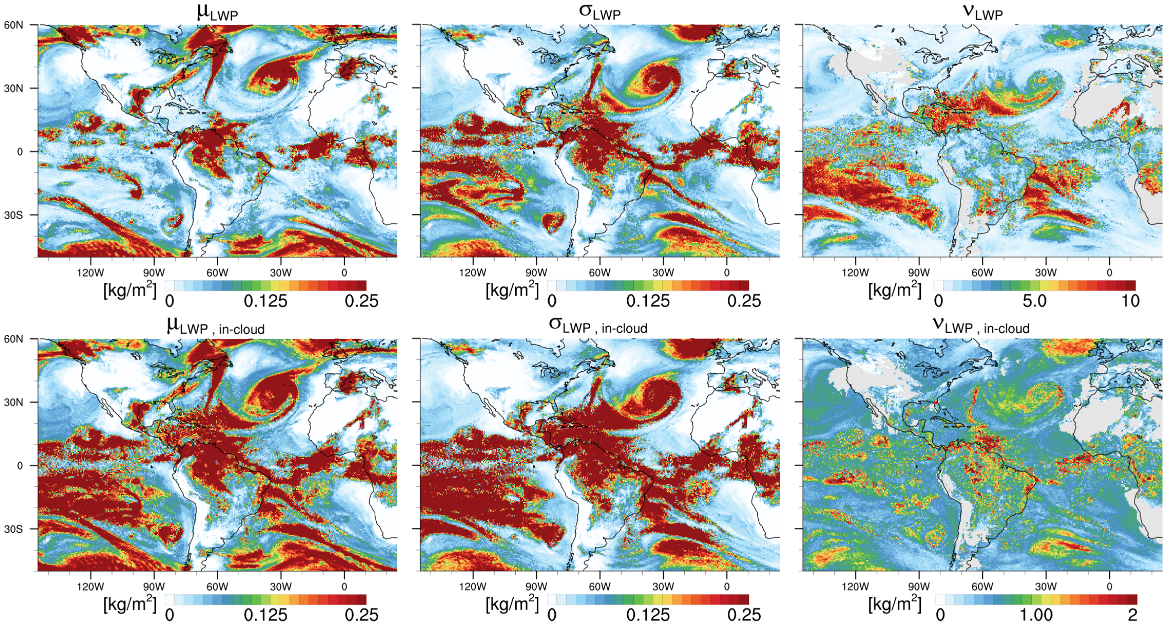


Figure 3.5.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: vertically integrated liquid water path (LWP) characteristics: Top row: LWP-PDF moments relative to grid-box scale with the LWP mean μ_{LWP} (left), standard deviation σ_{LWP} (middle), dispersion ν_{LWP} (right). Bottom row: LWP-PDF moments relative to in-cloud scale with the LWP mean μ_{LWP_c} (left), standard deviation σ_{LWP_c} (middle), dispersion ν_{LWP_c} (right). Grey color indicates cloud free regions, i.e. missing values.

3.2.1. Simple diagnostic schemes in CLUBB

The performance of the simple diagnostic schemes (uniform, triangular, and single Gaussian) introduced in Chapter 2 are now explored in the CLUBB framework. For this purpose they are diagnostically coupled to the grid-box saturation deficit PDF based on μ_s and σ_s. This approach is applied off-line, hence the resulting cloud properties do not feed back into the simulation which allows a consistent comparison. For their comparison the differences relative to the uniform box PDF are computed (Fig. 3.6), which is beneficial as the box PDF represents the most simplistic distribution function, apart from a delta function which can be considered as an all-or-nothing cloud scheme. The results for the vertically integrated cloud layer are shown with no loss in generality, as the individual cloud layers show a coherent signal.

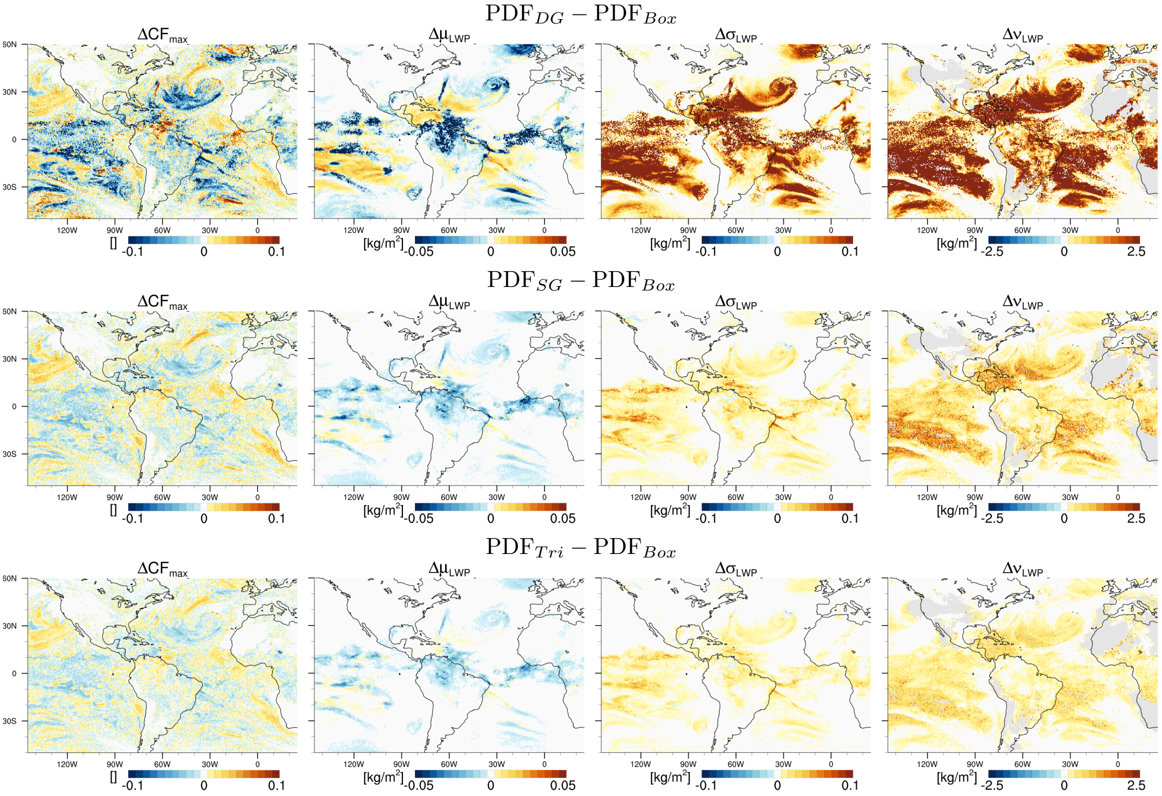


Figure 3.6.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: anomalies of vertically integrated liquid water path (LWP) characteristics, i.e. difference between the original double Gaussian (DG) and the uniform box PDF (top row), difference between the single Gaussian (SG) and the box PDF (middle row) and difference between the triangular (Tri) and box PDF (bottom row). Column 1: cloud fraction difference. Column 2: LWP mean difference $\Delta\mu_{\mathrm{LWP}}$. Column 3: LWP standard deviation difference $\Delta\sigma_{\mathrm{LWP}}$. Column 4: LWP dispersion difference $\Delta\nu_{\mathrm{LWP}}$. Grey color indicates cloud free regions, i.e. missing values.

Fig. 3.6 highlights that generally the differences to the box PDF increases with the complexity of the assumed PDF. In this connection the anomalies follow a coherent pattern, i.e. the sign of the anomalies in a particular region is the same. The unimodal distributions are more similar to each other than to the original CLUBB scheme, which highlights the importance of the second mode that is realized by the double Gaussian.

Cloud cover (CF_{max}) exhibits positive and negative anomalies. Positive anomalies indicate that CF_{max} of the box PDF is smaller. For the stratocumulus region a general positive bias occurs, which implies that CF_{max} of the box PDF is smaller. The difference is of the order of 3% for the triangular, 4% for the single Gaussian and 5% for the double Gaussian PDF. Accompanying errors in μ_{LWP}, σ_{LWP} and ν_{LWP} are marginal. One can conclude that the performance of the uniform PDF is remarkably good which is facilitated by the overall small skewness and similar the behavior of the PDF_1 and PDF_2.

In shallow cumulus regions cloud cover tends to be slightly overestimated by the box PDF. While the differences, with respect to the triangular and single Gaussian PDF, are similarly in the order of a few percent the difference to the double Gaussian PDF is larger and can reach 10%. Conceptually the alternating behavior for boundary layer clouds with small cloudiness can be understood from the theoretical overview (Fig. 2.2): For the same subgrid-scale variability and $\text{CF} < 0.5$ the box PDF has a more efficient shape to place its values in the saturated part. The higher the subgrid-scale variability is, the more pronounced the cloud properties differences will be. This gets evident further along the trade-wind trajectory as the skewness of the double Gaussian increases. Outlined in the previous section, parts of the shallow cumulus region exhibit extreme liquid water variability that results from PDF_1 being very wide. Since the relative weight of PDF_1 in these cases is very small, the simple diagnostic PDFs are merely affected. The notably big advantage is, in terms of a more realistic σ_{LWP}, of slanted PDF tails which is featured already by the slightly more complex triangular PDF. One can recognize the strong overestimation of μ_{LWP} produced by the box PDF within deep convective clusters, which is a feature somewhat independent from cloud cover and LWP variability trends. The negative anomalies arise mostly from levels above the boundary layer. Analyzing the contribution from individual levels reveals that cloud cover and μ_{q_l} of the box PDF are persistently larger than the double Gaussian.

Finally a notable point is that σ_{LWP} is mostly and ν_{LWP} is always positively biased, which follows the conceptual understanding that the box PDF is most restricted in producing liquid water variability and the associated dispersion is smallest.

3.3. ICON EDMF-DualM

One essential difference between the CLUBB and EDMF-DualM is the different allocation of physical meaning to the area partitioning of the double Gaussian PDF. In CLUBB the area fractions can vary smoothly between 0 and 1, while in EDMF-DualM the area fraction of PDF_1 is limited to 0 and 0.1 and PDF_2 between 0.9 to 1. The key idea in EDMF-DualM is that PDF_1 reflects the moist convective updraft PDF, while PDF_2 is the turbulent diffusive one. Additionally, within the mixed layer the area fractions are fixed

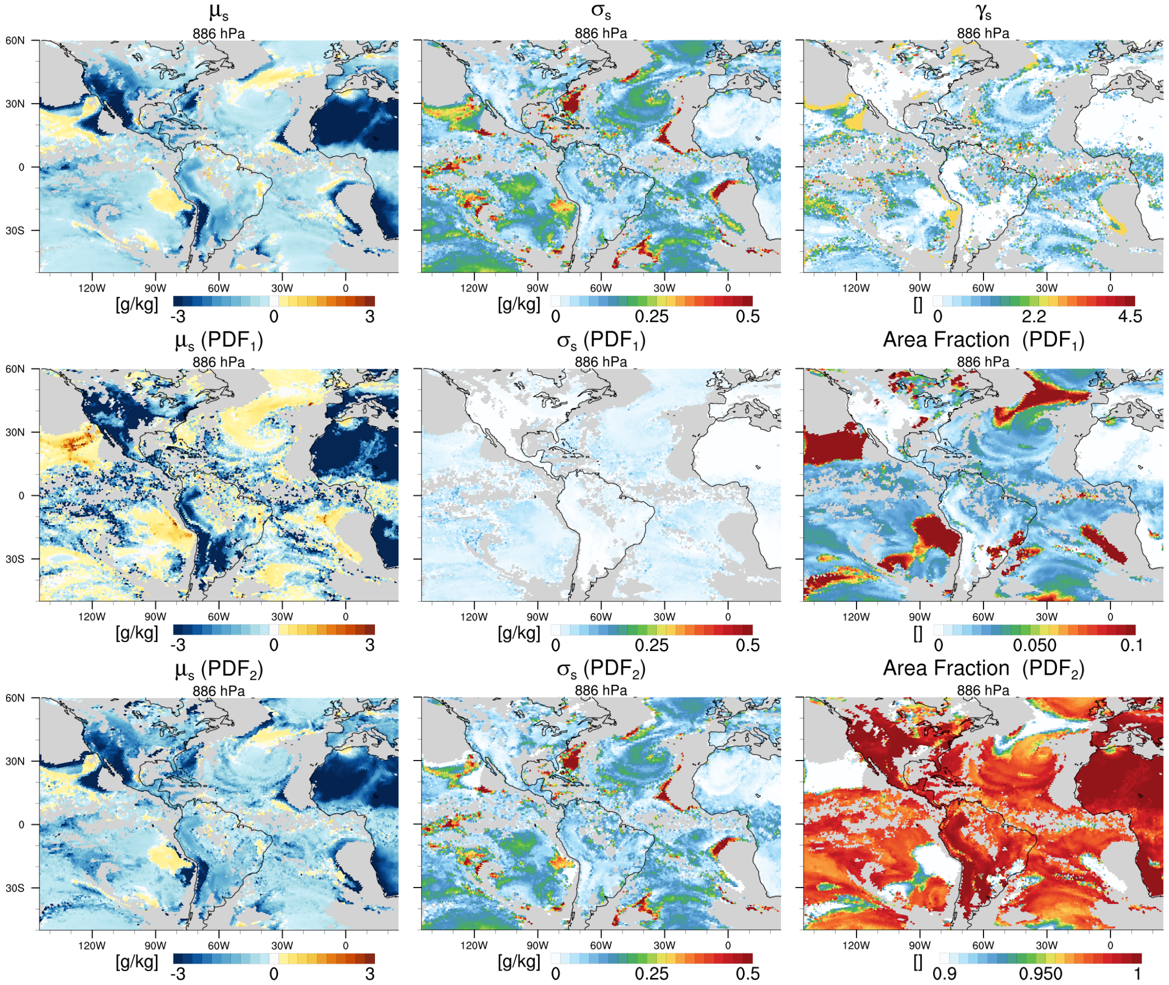


Figure 3.7.: Daily average based on 3 hourly output of the second day of a forecast simulation with the ICON-GCM model using the EDMF-DualM cloud parameterization showing the characteristics of the double Gaussian saturation deficit (s) PDF predicted by the EDMF-DualM. Grid-boxes above the EDMF-DualM layer are set to missing value (gray). Top row: grid-box mean μ_s (left), standard deviation σ_s (middle) and skewness γ_s (right). Middle row: contribution from PDF_1 and area fraction, which is the relative weight of this PDF component (right). Bottom row: contribution from PDF_2. Note that the color-range of the area fractions differ to CLUBB (Fig. 3.1). The model level that exhibits the highest stratocumulus cloud fraction is shown. The corresponding average marine subtropical pressure height is 886 hPa.

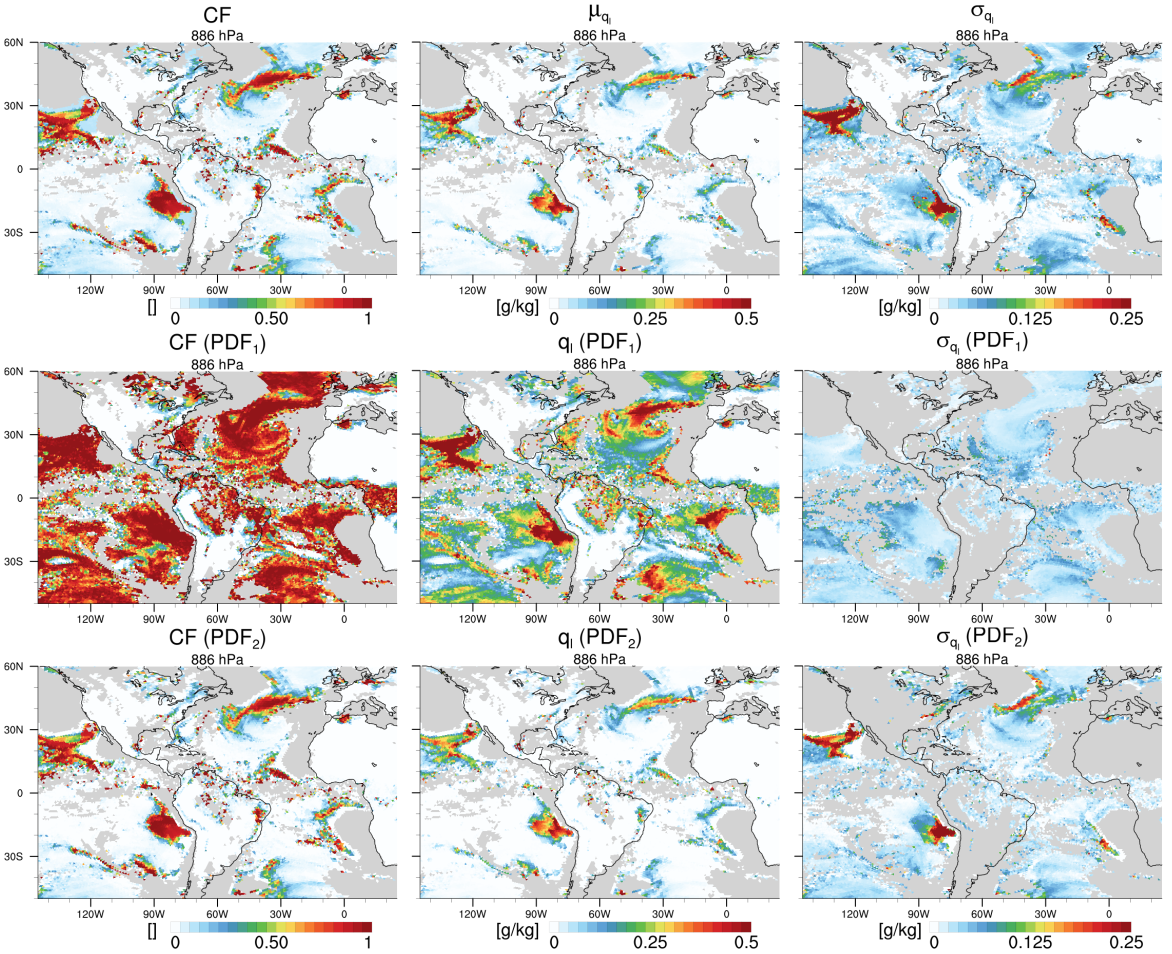


Figure 3.8.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: characteristics of the double Gaussian cloud PDF. Top row: grid-box cloud fraction CF (left), liquid water mean μ_{q_l} (middle) and standard deviation σ_{q_l} (right). Middle row: contribution from PDF$_1$. Bottom row: contribution from PDF$_2$. Grid-boxes above the EDMF-DualM layer are set to missing value (gray).

to a constant value with height, while in CLUBB each grid-box area fraction is diagnosed from the corresponding skewness.

Analogously to the previous section the mutual relationships of s-PDF and q_l-PDF as well as the vertically integrated LWP-PDF statistics produced by the EDMF-DualM cloud parameterization are explored. In order to preserve cloud regimes of a given synoptic situation, daily averages based on 3 hourly output of the second day of a forecast simulation with the ICON-GCM model are analyzed. To demonstrate the spatial distribution of the s-PDF (Fig. 3.7), the q_l-PDF (Fig. 3.8) and the grid-scale and in-cloud scale q_l statistics (Fig. 3.9) are shown on the 866 hPa level, which exhibits the highest stratocumulus cloud fraction. In conjunction, the vertically integrated statistics of cloud properties (Fig. 3.10) and liquid water path LWP-PDF statistics (Fig. 3.11) are presented. The vertically integrated is derived for the full vertical extent of the EDMF-DualM layer, which is basically

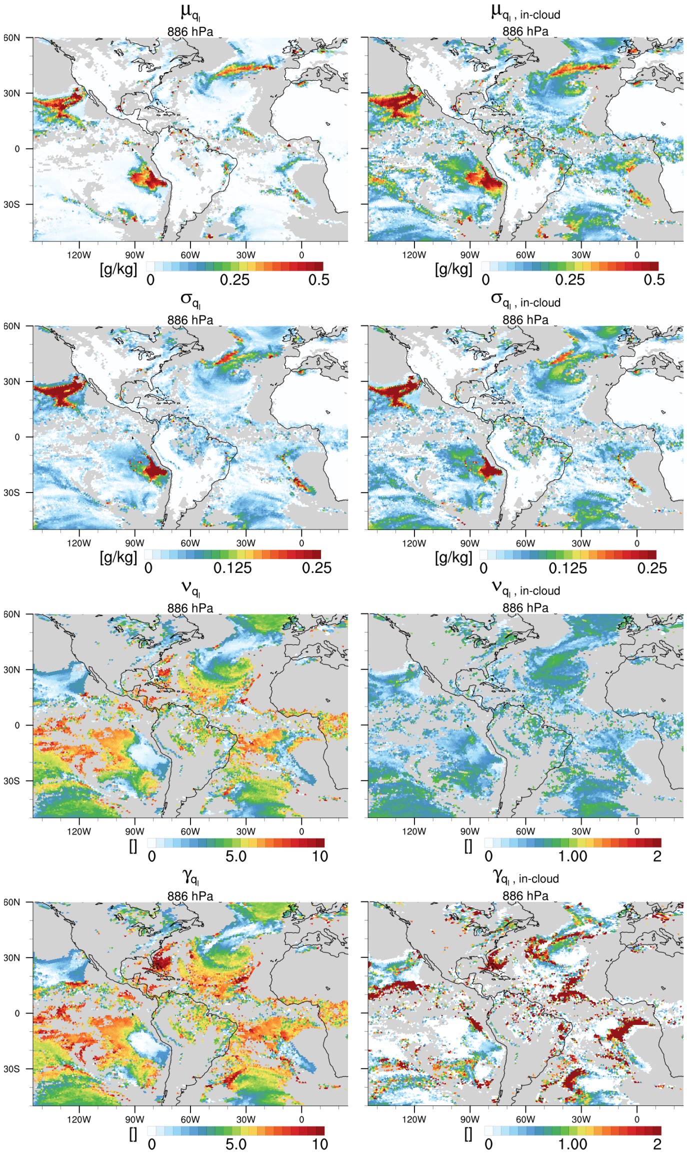


Figure 3.9.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: grid-box mean liquid water PDF statistics on grid-box scale (left) and on in-cloud scale (right). Liquid water: mean μ_{q_l} (row 1), standard deviation σ_{q_l} (row 2), dispersion $\nu_{q_l} = \sigma_{q_l}/\mu_{q_l}$ (row 3), skewness σ_{q_l} (row 4). Grid-boxes above the EDMF-DualM layer are set to missing value (gray).

limited to about 700 hPa (dry convection plumes might reach higher over deserts).

Two major properties of the general EDMF-DualM scheme behavior are revealed, which are inherently different to the CLUBB scheme: a) the temporal frequency of occurrence of scheme (Fig. 3.10, top-right). If the test-parcel ascent indicates that no boundary layer convection is present within a column, EDMF-DualM is passive and no s-PDF and associated cloud properties are produced. b) the vertical extent of the EDMF-DualM boundary layer which is evaluated for each time step based on the height reached by

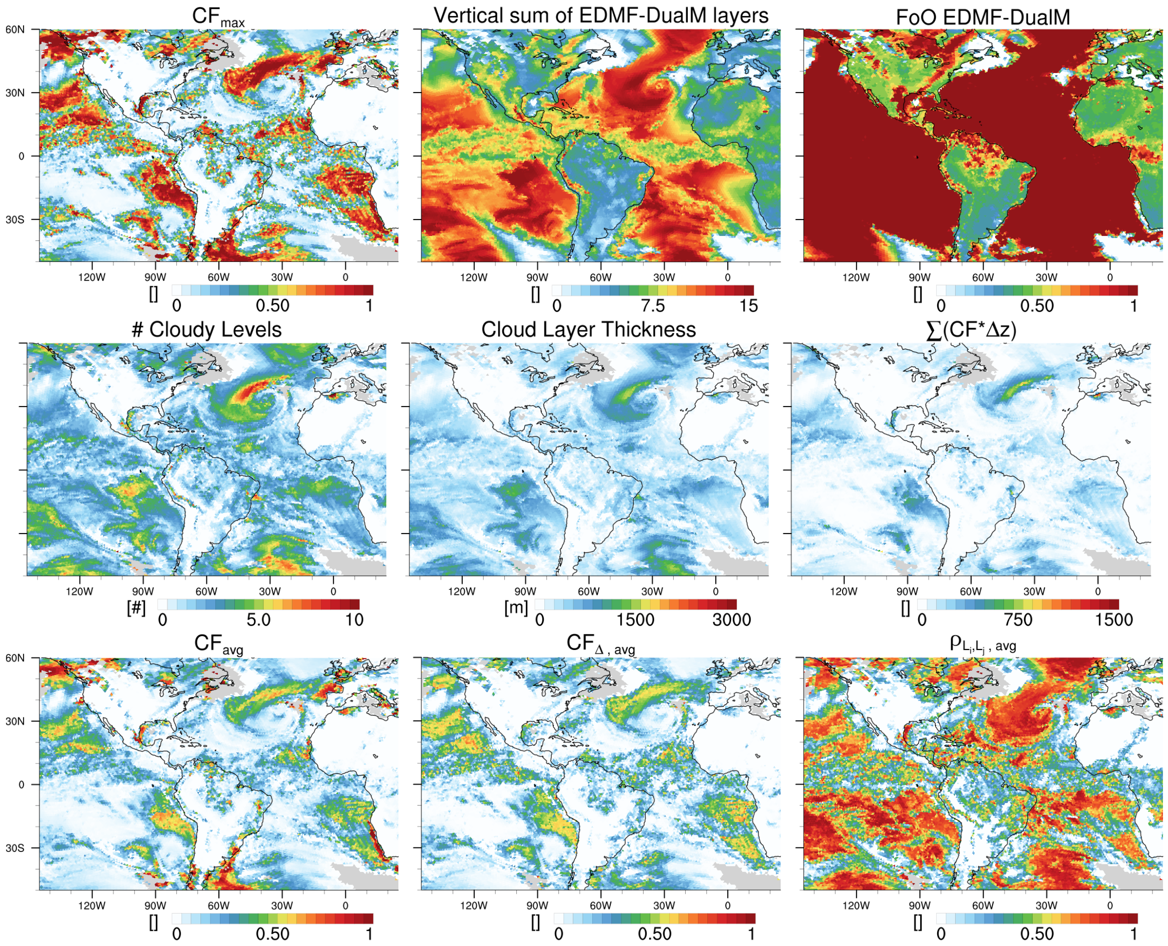


Figure 3.10.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: vertically integrated cloud characteristics. Top row: maximum cloud fraction (CF) within a column (left), average column sum of EDMF-DualM layers (middle), temporal frequency of occurrence of EDMF-DualM activity (right). Middle row: number of cloud layers within a column (left), cloud layer thickness (middle), sum of the layer product of CF times layer thickness (right). Bottom row: average column CF (left), average inter-layer CF difference (middle), bulk inter-layer correlation coefficient (right). Columns with no EDMF-DualM activity are set to missing value (gray).

the test parcel (Fig. 3.10, top-middle). This approach marks an important conceptual difference to CLUBB, which defines the scope of the EDMF-DualM scheme solely on the unified treatment of the boundary layer transport, but leaving deep convective transport out. c) while the s mean of each PDF component is quite similar, the variability of each PDF nearly exhibits an opposite pattern: in contrast to CLUBB in EDMF σ_s of PDF$_1$ has moderate values and is relatively homogeneously distributed, PDF$_2$ exhibits a rich spatial variability and occasionally values one order in magnitude larger. Nevertheless the trend of increasing σ_s with increasing convective activity is similar.

Marine stratocumulus

For marine stratocumulus the area fraction partitioning has a distinct state of $\mathrm{PDF}_1 = 0.1$ and $\mathrm{PDF}_2 = 0.9$, which results from a cloud regime specific case switch. The saturation deficit means of the individual Gaussian PDFs exhibit a similar patter compared to CLUBB, but there are two important differences (Fig. 3.7): a) the magnitude of supersaturation is higher for PDF_1, which however influences the grid-box mean s only slightly as the area fraction of PDF_1 is much smaller than in CLUBB. b) The spatial extent of supersaturated regions is higher, even though the chosen ICON model layer is already higher up. c) There is a larger difference in the spatial pattern between PDF_1 and PDF_2. While a PDF_2 grid-box can be already above the inversion, the convective PDF_1 is still supersaturated, indicating a small scale penetration of the free troposphere.

PDF_2 has more extreme σ_s at the edge of stratocumulus regions: when tracking the development of these zones across vertical levels this feature is propagating eastward with high, thus reflecting the slanted in marine boundary height increase along the trade-wind trajectory. The high σ_s along this slope is a consequence of intense turbulent variance production which is driven by strong vertical humidity gradients. Still, with respect to the cloud properties this zone has a minor importance because μ_s of this PDF_2 low (Fig. 3.7). CF of PDF_1 generally close to 1 (for shallow cumulus too), which results from the slightly supersaturated μ_s alongside with little variance so that the whole PDF_1 fits in the saturated part (Fig. 3.8). In particular regions of larger μ_s exhibit high μ_{q_l}, which however do not overly contribute to the grid-box mean as area fraction is small. In stratocumulus grid-box and in-cloud scale μ_{q_l} and σ_s deviate significantly form CLUBB and are much larger (higher than in other regions), while $\nu_{\mathrm{LWP}}$and γ_{LWP} are comparable. This pattern is essentially equivalent to the vertically integrated LWP-PDF statistics since the number of cloudy model layers is small (Fig. 3.11).

Marine stratocumulus transition

The transition from closed cell stratocumulus to shallow cumulus is rather abrupt in terms of PDF area partitioning, i.e. the area fraction of PDF_1 reduces from 0.1 to 0.03, which clearly reveals the internally used convection types. In contrast, the transition in terms of cloud properties is much smoother, which is shown by the gradually increasing number of cloud layers or cloud thickness (Fig. 3.10). In particular the open cell stratocumulus in the south-east Pacific evolve into a quite vigorous cloud deck that is supported by a supersaturated PDF_1. Similar to the closed cell stratocumulus point c), the open cell stratocumulus of the cold air outbreak in the North Atlantic exhibits a supersaturated PDF_1 while PDF_2 is slightly negative (response to the strong surface heating). The open cell stratocumulus with medium to high cloud cover marks the most intense cloud type that is produced by the EDMF-DualM and is in many respects similar to the analog regime

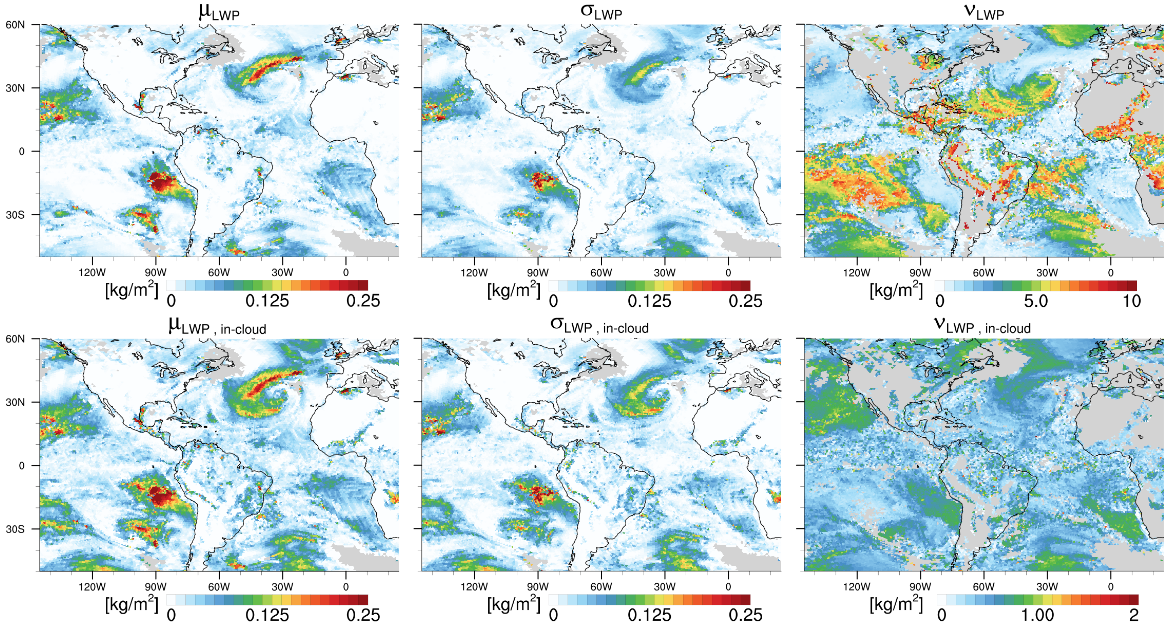


Figure 3.11.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: vertically integrated liquid water path (LWP) characteristics: Top row: LWP-PDF moments relative to grid-box scale with the LWP mean μ_{LWP} (left), standard deviation σ_{LWP} (middle), dispersion ν_{LWP} (right). Bottom row: LWP-PDF moments relative to in-cloud scale with the LWP mean μ_{LWP_c} (left), standard deviation σ_{LWP_c} (middle), dispersion ν_{LWP_c} (right). Columns with no EDMF-DualM activity are set to missing value (gray).

in CLUBB. While in CLUBB this region frequently transitions to even deeper convection one might wonder what would be happen if the EDMF-DualM would not be restricted to the boundary layer.

Marine shallow cumulus

For the marine shallow cumulus case the area partition of the PDF_1 varies smoothly between 0 and 3%. These values are motivated by the typical area fractions covered by organized updrafts of prototype marine shallow cumuli. μ_s of the updraft PDF_1 is close to saturation and its variability is small, which reflects the idea of an relative undiluted moist convective updraft mode. μ_s of the diffusive PDF_2 is moderately negative and its variability is relatively large. While grid-box cloud fraction and μ_{ql} is mainly supported by the updraft PDF, σ_{ql} gets contributions from both PDFs, in particular the long positive tail of the diffusive PDF is extending into the saturated range. Considering the cloud properties on the in-cloud scale the importance of the updraft PDF gets emphasized. The in-cloud liquid water skewness is small which is a result from the almost fully saturated updraft PDF_1, which than resembles the underlying single Gaussian PDF that is symmetric. In comparison to the single layer, the vertical integral highlights the cloud characteristics of the slanted marine boundary layer development along the trade-wind trajectories.

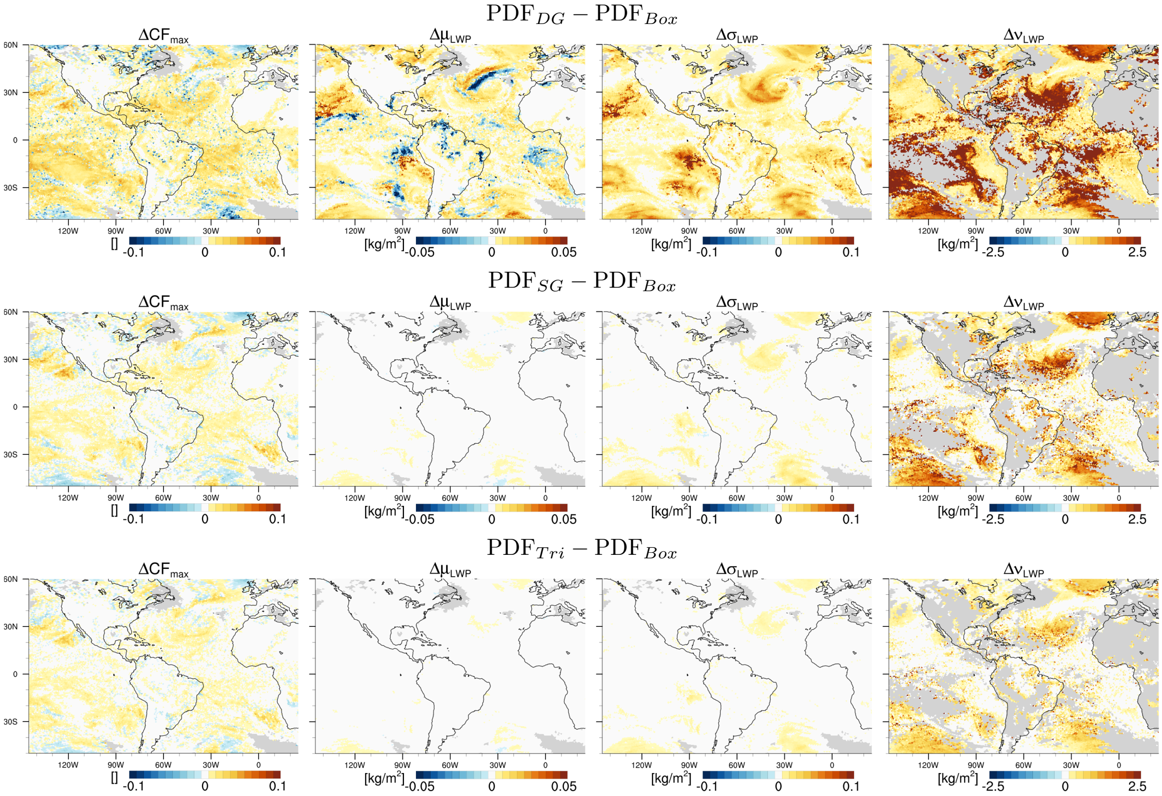


Figure 3.12.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: anomalies of vertically integrated liquid water path (LWP) characteristics, i.e. difference between the original double Gaussian (DG) and the uniform (box) PDF (top row), difference between the single Gaussian (SG) and the box PDF (middle row) and difference between the triangular (Tri) and box PDF (bottom row). Column 1: cloud fraction difference. Column 2: LWP mean difference $\Delta\mu_{\mathrm{LWP}}$. Column 3: LWP standard deviation difference $\Delta\sigma_{\mathrm{LWP}}$. Column 4: LWP dispersion difference $\Delta\nu_{\mathrm{LWP}}$. Columns with no EDMF-DualM activity are set to missing value (gray).

Cumulus over land

EMDF-DualM shows much weaker temporal activity over land, which is primarily triggered by the response of solar heating on different surface properties in relation to the diurnal cycle. Deeper convective regions connected to the ITCZ are captured but shallow convective regions exhibit underrepresented cloud properties (Chapter 5 investigates this point further).

3.3.1. Simple diagnostic schemes in EDMF

Analog to CLUBB, the performance of the simple diagnostic schemes (uniform, triangular, and single Gaussian) introduced in Chapter 2, are now explored in the EDMF-DualM framework. Comparing the behavior of the simple diagnostic schemes coupled to CLUBB (Fig. 3.6) with the behavior of the simple diagnostic schemes coupled to EDMF-DualM reveals essentially a coherent signal. Even though the same diagnostic closures are used, this characteristic is not necessarily expectable because different combinations of μ_s and σ_s are used and the different vertical staggering options are possible.

Cloud cover anomalies are in the same range as in Fig. 3.6, however the bias is largely positive because the negative anomalies introduced by deeper clouds in CLUBB are not simulated by EDMF-DualM. Still, similar to CLUBB, the box PDF produces negative μ_{LWP} anomalies in regions of the thickest EDMF-DualM clouds. μ_{LWP} and σ_{LWP} anomalies among unimodal PDFs are very small, which underlines the perception that the difference between unimodal and bimodal PDFs are more important than the shape within each mode.

3.4. Vertical histograms

To gain more insight about the vertical variation of cloud properties, their joint histograms with height are computed (Figs. 3.13, 3.14). Using a joint histogram has some distinct advantages compared to a average vertical profile. First the properties might exhibit a multi-modal distribution, an information that would be lost by considering just the layer mean and secondly one avoids the averaging problem introduced by cloud-free and missing grid-boxes. For the construction of the vertical joint histograms the instantaneous model output over the 145°W to 25°E, 50°S to 60°N domain is used and only cloudy grid-boxes are considered. By construction the frequency of occurrence (FoO) in each sub-figure adds up to 100%. The results from the original CLUBB and EDMF-DualM cloud scheme are shown, as well as for the respective diagnostic single Gaussian PDF. Generally one can note the higher vertical resolution of ICON (20 levels) relative to the GFDL-AM3 (13 levels) below 650 hPa.

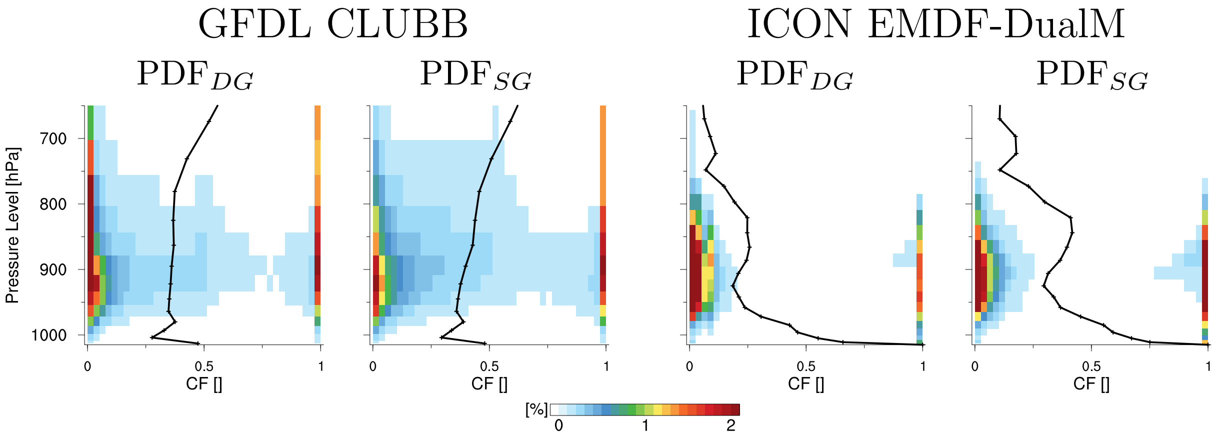


Figure 3.13.: Vertical histograms of cloud fraction (CF) from GFDL-AM3 CLUBB and ICON EDMF-DualM instantaneous 3D model output, setup equivalent to Figs. 3.1 and 3.7. Original double Gaussian PDF (DG) of the CLUBB scheme (column 1) and the corresponding diagnostically coupled single Gaussian (SG) (column 2), while (column 3) shows the original DG of the EDMF-DualM scheme and the corresponding SG (column 4).

The vertical distribution of cloud fraction (CF) is shown in Figs. 3.13, 3.14 which highlights the different core ideas of both schemes. CLUBB incorporates deep convection and is therefore vertically unbounded, while EDMF-DualM activity is fading out above 700 hPa. Both schemes favor extreme CF values, while CLUBB show at least occasionally intermediate CF values, while they practically do not occur in EDMF-DualM. Considering only cloudy grid-boxes it is notable that the diagnostic single Gaussian shows slightly higher CF on average.

In-cloud liquid water increases with height with similar distributions, except the EDMF-DualM double Gaussian shows an isolated mode at 0.25 g/kg, which is introduced by the moist convective updraft of PDF_1.

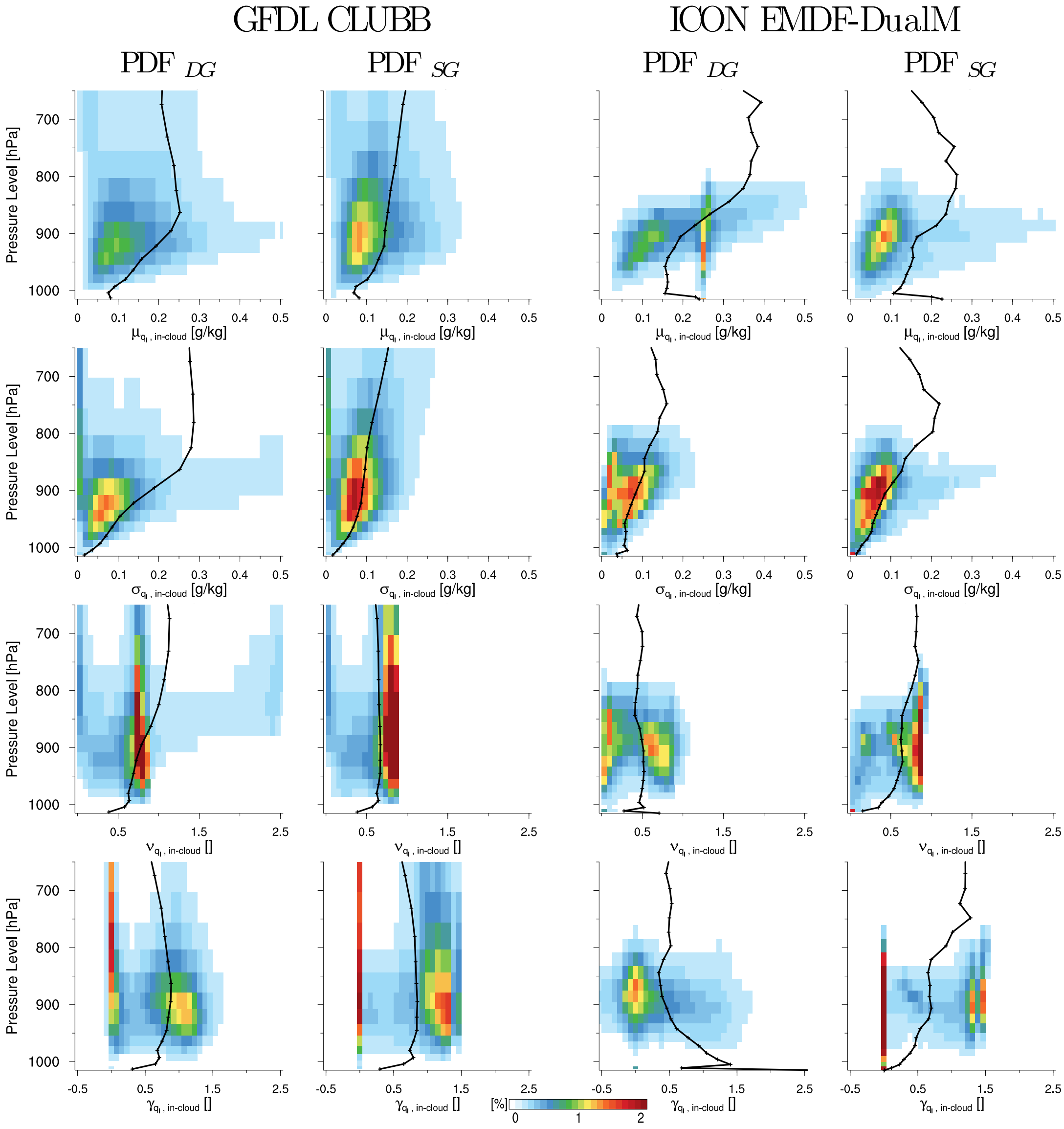


Figure 3.14.: Vertical histograms analog to Fig. 3.13, now for the in-cloud scale of liquid water $q_{l,c}$-PDF. Row 1: $q_{l,c}$-mean ($\mu_{q_{l,c}}$). Row 2: $q_{l,c}$-standard deviation ($\sigma_{q_{l,c}}$). Row 3: $q_{l,c}$-dispersion ($\nu_{q_{l,c}}$). Row 4: $q_{l,c}$-skewness ($\gamma_{q_{l,c}}$).

Ther vertical pattern of $\sigma_{q_{l,c}}$ is similar to $\mu_{q_{l,c}}$. The moist convective updraft PDF$_1$ is less distinct and exhibits very low $\sigma_{q_{l,c}}$ (undiluted). $\sigma_{q_{l,c}}$ of CLUBB shows the extremely large values that are associated to the PDF$_1$ in shallow cumulus regions along with low CF.

Dispersion $\nu_{q_{l,c}}$ of a single Gaussian is bounded between 0 and 0.7 (Chapter 2), which is reflected by the diagnostically coupled ones. Theoretically unbounded positive dispersion could be created by the double Gaussian PDFs, however they exhibit a very similar distribution than their diagnostic single Gaussians and only occasionally feature slightly higher values. This result has important implications for a potential coupling of microphysics and radiation parameterizations. A noteworthy exception occurs for the double Gaussian of CLUBB showing another weak local maximum at 2.5.

Similar to $\nu_{q_{l,c}}$ the in-cloud liquid water skewness of single Gaussian PDF is bounded between 0 and 1.5. Notably is the quite different behavior among the double Gaussian PDFs, which was not as evident for the other cloud parameters. Both double Gaussian PDFs are basically limited to $\nu_{q_{l,c}}$ of 1.5 too, but as outlined in the previous section, CLUBB uses a partially saturated single Gaussian with a large area fraction, while EDMF-DualM exhibits a fully saturated single Gaussian with a small area fraction. Additionally negative in-cloud liquid water skewness is produced by the EDMF-DualM. This scenario occurs in the shallow cumulus regime, when the updraft is fully saturated and behaves like a delta function and the tail of the diffusive PDF is cloudy too while μ_{q_l} of PDF_1 is larger than μ_{q_l} of PDF_2.

3.5. Joint histogram with cloud fraction

In this section the mutual relationships between CF vs. s-PDF and CF vs. q_l-PDF are further explored for the CLUBB and the EDMF-DualM scheme. Binning the cloud properties in terms of the CF allow to decouple the influence of CF as theoretically explored in Chapter 2. The joint histograms are computed using 25 bins in each dimension and normalizing the frequency of occurrence to 100 %. As outlined in Chapter 2, cloud fraction is deeply connected to the normalized shape of unimodal PDFs.

$\mu_{q_{l,c}}$ is high at low CF for the double Gaussian in CLUBB and EDMF-DualM as well as for the diagnostic single Gaussian in EDMF-DualM. With slightly increasing CF the majority in of $\mu_{q_{l,c}}$ is low and subsequently increases with CF. The peak in $\mu_{q_{l,c}}$ is connected to the shallow cumulus regime and is realized by the PDF_1.

$\sigma_{q_{l,c}}$ follows the same pattern as $\mu_{q_{l,c}}$. The distinct characteristics of the EDMF-DualM updraft PDF, bounded between $0 < \mathrm{CF} < 0.1$, are highlighted.

The in-cloud dispersion summarizes their mutual effect and highlights the distinct relationship of the single Gaussian PDF and points out two major specialties of CLUBB and EDMF-DualM. First, the high $\sigma_{q_{l,c}}$ shallow cumulus mode that became evident in Fig. 3.4 shows a distinct second branch in the CF-dispersion phase-space, while the majority follows the single Gaussian relationship. Second, EDMF-DualM follows primarily the single Gaussian relationship too, but the very small variance at low CF creates a lower $\nu_{q_{l,c}}$ that is characteristic for a uniform PDF, or more extreme a delta function. Additionally at $\mathrm{CF} = 1$ larger $\nu_{q_{l,c}}$ is possible.

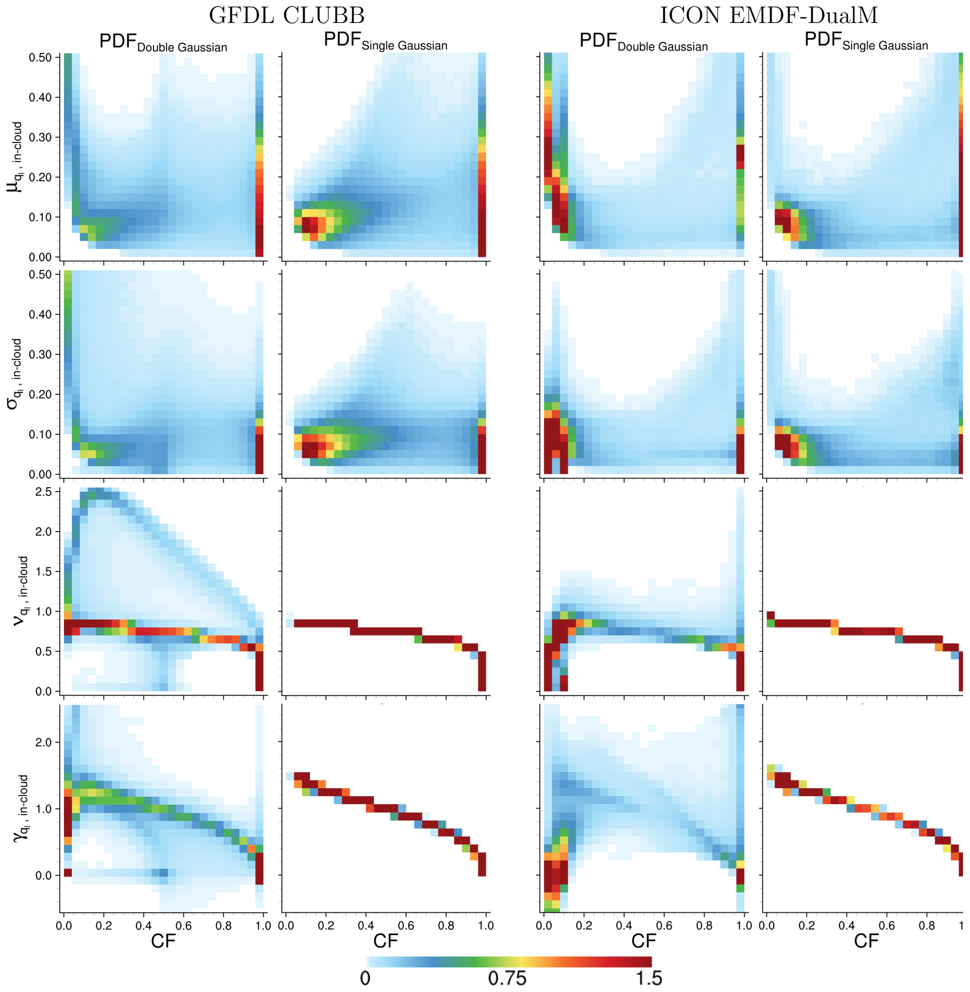


Figure 3.15.: Joint histograms of cloud fraction (CF) vs. in-cloud liquid water $q_{l,c}$-PDF moments analog to Fig. 3.13. Row 1: $q_{l,c}$-mean ($\mu_{q_{l,c}}$). Row 2: $q_{l,c}$-standard deviation ($\sigma_{q_{l,c}}$). Row 3: $q_{l,c}$-dispersion ($\nu_{q_{l,c}}$). Row 4: $q_{l,c}$-skewness ($\gamma_{q_{l,c}}$).

3.6. Conclusion

The characteristics of subgrid-scale saturation deficit and related cloud properties, produced by the recently developed unified boundary layer transport and convection schemes 'CLUBB' and 'EDMF-DualM', are explored using short synoptically driven GCM forecast simulations. CLUBB, the unified higher-order turbulence transport- and cloud parameterization (Golaz et al., 2002a; Larson and Golaz, 2005; Larson et al., 2012), is used within the GFDL-AM3 model (Guo et al., 2014) along with the newly development capability to parameterize deep convection as well (Guo et al., 2015). EDMF-DualM, the eddy-diffusivity mass flux scheme extended to a dry and a moist convective mass flux (Neggers et al., 2009), is used within the new ICON (ICOsahedral Non-hydrostatic) general circulation model (Zängl et al., 2015). Both schemes incorporate a statistical cloud parameterization which use of a double Gaussian PDF to predict cloud cover and liquid water. Even though these parameterizations are quite encapsulated in both models, a method to diagnose all related PDF parameters was developed. This was achieved in a consistent manner which provides the opportunity to reproduce the mutual relationships between the s-PDF and the cloud properties. Using the grid-box mean properties of the envelope s-PDF, simpler unimodal cloud closure were coupled to both models diagnostically. In order to compare the bulk effect of the parameterizations behavior and in preparation for their evaluation in Chapter 5, the analytical vertical integration method (developed in the previous Chapter 2) for subgrid cloud water variability was applied.

Both schemes produce extensive stratocumulus cloud decks prevailing over the eastern subtropical oceans. The high cloud cover mainly originates from a single model level. Additionally both schemes feature a transition from marine closed cell stratocumulus (high cloud cover) to open cellular convection (medium cloud cover) to shallow cumulus convection (low cloud). The transition is accomplished in the double Gaussian PDF framework by μ_s of PDF_1 remaining weakly positive while its area fraction decreases with increasing convective activity, along with μ_s of PDF_2 being moderately negative with small μ_s, but produces small cloud amounts with little water variability. The representation of shallow cumulus exhibits a conceptual difference between the CLUBB and EDMF-DualM because they allocate a different physical meaning to the double Gaussian PDFs, which results in a different area partitioning. Shallow cumulus is represented in CLUBB by a partially saturated single Gaussian with large area fraction, while EDMF-DualM exhibits a fully saturated single Gaussian with a small area fraction. Even though both scenarios result in the similar cloud fractions, the associated cloud properties are different, in particular the in-cloud liquid water skewness show a different behavior. In CLUBB shallow cumulus is frequently developing into deeper convection, while EDMF-DualM is limited to the boundary layer. EDMF-DualM shows a strong land-sea contrast in its temporal frequency of occurrence, i.e. activity.

Both schemes favor extreme cloud fraction. However, while CLUBB exhibits intermediate CF occasionally, they practically do not occur in EDMF-DualM. Liquid water dispersion, i.e. the ratio of variability divided by its mean, shows a more coherent signature with larger $\nu_{q_{l,c}}$ or ν_{LWP_c} values in the cumulus regions than stratocumulus areas.

The behavior of the simple diagnostic schemes (uniform, triangular, and single Gaussian), that were introduced in Chapter 2, is explored relative to the double Gaussian PDFs of CLUBB and EDMF-DualM. The comparisons are performed off-line, therefore the resulting cloud properties do not feed back into the simulations, which allows a consistent comparison. Generally, the differences relative to a uniform box PDF are increasing with complexity of the assumed PDF and the anomalies follow a coherent pattern, i.e. the sign of the anomalies in a particular region is the same. The unimodal distributions are more similar among each other than to the double Gaussian PDFs. In particular the difference gets larger, the more CLUBB and EDMF-DualM make use of their two Gaussian PDFs, i.e. the more bimodal they are. One can conclude that the difference between unimodal and bimodal PDFs is more important than the shape within each mode. The characteristics of the simple unimodal PDFs diagnosed from CLUBB and EDMF-DualM are quite similar, which is not necessarily expected because different input combinations of μ_s and σ_s and different vertical staggering permutations are possible. Cloud cover and liquid water path anomalies can be positive and negative, while liquid water path variability and dispersion anomalies are generally positive, i.e. larger compared to the box PDF. In stratocumulus regions positive cloud cover anomalies occur. Vice versa for shallow cumulus regions, the box PDF produces slightly higher cloud cover, which can conceptually be understood from its shape. Once saturation occurs, a shape with no slanted PDF tails is more efficient to place its values in the saturated part for the same subgrid-scale variability.

Using joint histograms of cloud fraction versus cloud properties, the important role of CF can be highlighted. CF is deeply connected to the normalized shape of unimodal PDFs. The in-cloud dispersion summarizes their mutual effect and highlights the distinct relationship of the single Gaussian PDF and points out two major specifics of CLUBB and EDMF-DualM. First, the high $\sigma_{q_{l,c}}$ shallow cumulus mode that became evident in Fig. 3.4 shows a distinct second branch in the CF-dispersion phase-space, while the majority follows the single Gaussian relationship. Second, EDMF-DualM follows primarily the single Gaussian relationship, but the very small variance at low CF creates lower $\nu_{q_{l,c}}$ that are characteristic for a uniform PDF.

Chapter 4

Importance of subgrid temperature variability for cloud parameterization

The question about the importance of subgrid-scale temperature variability for saturation deficit variability is motivated by the circumstance that most statistical cloud schemes are just formulated in terms of a humidity variability. Doing so they assume a constant saturation humidity across the grid-box, which is equivalent to neglecting temperature variability, leaving humidity fluctuations as the only predictor. Therefore the question arises if clouds are really so strongly humidity driven, or what the consequences of neglected temperature variability in terms of cloud properties are.

Saturation humidity is a function of temperature and pressure. Utilizing the Clausius-Clapeyron equation one can estimate that temperature fluctuations are several orders of magnitude more relevant (Bougeault, 1981a) than pressure fluctuations. Therefore the following discussion will focus on the temperature variability. The importance of temperature variability for cloud diagnostics was addressed in early theoretical studies by Mellor (1977) and Sommeria and Deardorff (1977), which were further explored using observational data by Larson et al. (2002) and Tompkins (2003) and implemented into the CLUBB cloud parametrization by Golaz et al. (2002a;b) (see Chapter 3). Recently these ideas where successfully implemented in GCMs (Guo et al., 2014; Bogenschutz et al., 2013). Based on observational data, temperature fluctuations are considered less important than humidity fluctuations, due to their shorter time-scales, as temperature perturbation are quickly removed by buoyancy adjustment (see Tompkins (2003) and references therein). The contribution of humidity-, temperature- and their covariability fluctuations can be combined into saturation deficit fluctuations, a concept that allows to study the relative importance Mellor (1977); Sommeria and Deardorff (1977). Little is known about the relative importance of temperature fluctuations for different cloud regimes, so the first research question (1) will explore spatial and vertical aspects of saturation deficit variability contributions. In particular the special role humidity- and temperature covariability is explored in more detail. Subsequently the second research question (2) will explore how neglected temperature variability translates into CF and q_l biases. Price and Wood (2002) addressed this question using aircraft data for the subtropical boundary layer and found that the mean bias in cloud cover is around 6%. Especially warm boundary clouds seem to be sensitive to temperature fluctuations, thus the cloud fraction and liquid water biases could translate into a significant cloud radiative effect (see Chapter 1).

4.1. Data and Method

In this study the advanced boundary layer transport and cloud parametrization CLUBB (Chapter 3) is used to address the influence of subgrid-scale temperature and humidity variability on CF and q_l in a qualitative manner. The scope is to reveal the mechanisms between subgrid-scale variability and cloud properties. For this purpose a GCM simulation of one month is performed, which is long enough to average out the footprint of synoptic features, which occur on timescales of a few days (to that extent this chapter differs from the short forecast simulation of Chapter 3). The 30 day simulation for May is initialized from a NCEP nudged base run. The focus is on the subtropical marine regions and warm boundary layer clouds as they are most relevant for the raised questions and especially prone to the cloud-climate uncertainty. The original version of CLUBB (Chapter 3) is compared to a modified one, in which the variability of the temperature PDFs are set to zero. This is done diagnostically, so that the modified CF and q_l cannot feed back into the model simulation. Hence the two parameterization setups do not produce different atmospheric states, which would complicate the attribution of CF and q_l differences.

Additionally, results from large-eddy simulations (LES) are used. In cloud resolving simulations the humidity and temperature distributions are an emergent phenomena. UCLA LES model data (Stevens et al., 2005) of prototype cloud regimes are analyzed, which describe idealized conditions along the subtropical trade-wind trajectory. The four cases are: closed cell stratocumulus (S12), stratocumulus above shallow cumulus (S11) and shallow cumulus (S6) which are all based on the CFMIP/GASS Intercomparison of Large-Eddy and Single-Column Models project (CGILS) (Blossey et al., 2013) and deeper shallow cumulus (RICO) based on Seifert and Heus (2013). The simulation set-ups are summarized in Table 4.1. The domain size of the two stratocumulus cases is rather small and humidity and temperature variability might be underestimated. However the horizontal scales of meso-scale circulations in stratocumulus are small and the focus is not the quantify the absolute magnitude of variability, but rather to improve the understanding of conceptual mechanisms. For the RICO case the use of a larger domain size is necessary since individual cumuli can reach above 4 km and alongside with considerable meso-scale organization that exhibit spatial scales multiple times larger than the boundary layer depth.

Table 4.1.: LES simulation setup for the prototype cloud regimes.

Case	Cloud type	Domain	Resolution	Reference
S12	Stratocumulus	2.4×2.4 km^2	25 m	(Blossey et al., 2013)
S11	Stratocumulus over cumulus	4.8×4.8 km^2	50 m	(Blossey et al., 2013)
S6	Trade cumulus	9.6×9.6 km^2	100 m	(Blossey et al., 2013)
RICO	Trade cumulus	50×50 km^2	25 m	(Seifert and Heus, 2013)

4.2. Subgrid-scale humidity and temperature variability

Using the GFDL-AM3 CLUBB GCM one benefits from the unified treatment of deep convection, shallow convection, cloud macrophysics and the planetary boundary layer transport, which implies that all their subgrid-scale fluctuations, are captured by the same multivariate PDF. Using the saturation deficit concept the thermodynamic part of this multivariate PDF can be transformed, so that contribution to s and σ_s makes the humidity- and temperature influence comparable. Our motivation here is to advance the understanding of the parameterization's subgrid-scale behavior, synthesized by σ_s, while grid-box mean s results from the predictive equation of the model. Therefore the discussion will focus on σ_s. The spatial pattern of the s-PDF and the related cloud properties were introduced for a particular synoptic situation in the Figs. 3.1, 3.2, and is additionally shown for the 30 day period and for multiple height levels in Fig. C.11, which will be discussed here. Especially over the south-east subtropical Pacific a persistent cloud deck shows CF close to 100%. Despite to the moderate vertical resolution, which blurs vertical gradients, the stratocumulus clouds occurs in a narrow range of height levels. Moving westward across the ocean basins, along the trade-wind trajectory, CF decreases and in-cloud $q_{l,c}$ increases, indicating the transition to shallow cumulus. Together with this smooth transition the marine boundary layer deepens, cloud tops are ascending and σ_s increases with height, acting to create stronger subgrid-scale humidity excesses (Chapter 3). Highest values of σ_s are found in regions of deeper convection, in the vicinity of the ITCZ. Above clouds σ_s strongly reduces, which gets particularly visible in subsidence regions where clouds are limited to low levels.

Spatial pattern of saturation deficit variability contributions

Saturation deficit variability σ_s is the linear combination of total specific humidity fluctuations q_t' (term I) and liquid water potential temperature fluctuations θ_l' (term III), as well as their covariability $\overline{q_t'\theta_l'}$ (term II), defined in Eq. (2.5) in Chapter 2. Their relative contributions to σ_s are shown in Fig. 4.1. A strong land-marine contrast between the q_t' and θ_l' contributions can be detected in the lower part of the troposphere, where different surface properties and associated surface fluxes, are the dominant driver. For temporal averages the correlation of $\overline{q_t'\theta_l'}$ is almost exclusively negative, so that the contribution of the covariability (term II) is generally positive. Considering single time steps of instantaneous model output the covariability contribution close to the surface can occasionally be negative (positive $\overline{q_t'\theta_l'}$ correlation). The land-marine contrast appears as well for term II, with a lower σ_s contribution over land compared to ocean. However this contrast reduces with height in regions that have higher σ_s (Fig. C.11). Dry regions over land have little σ_s and temperature variability contributes most. The temperature

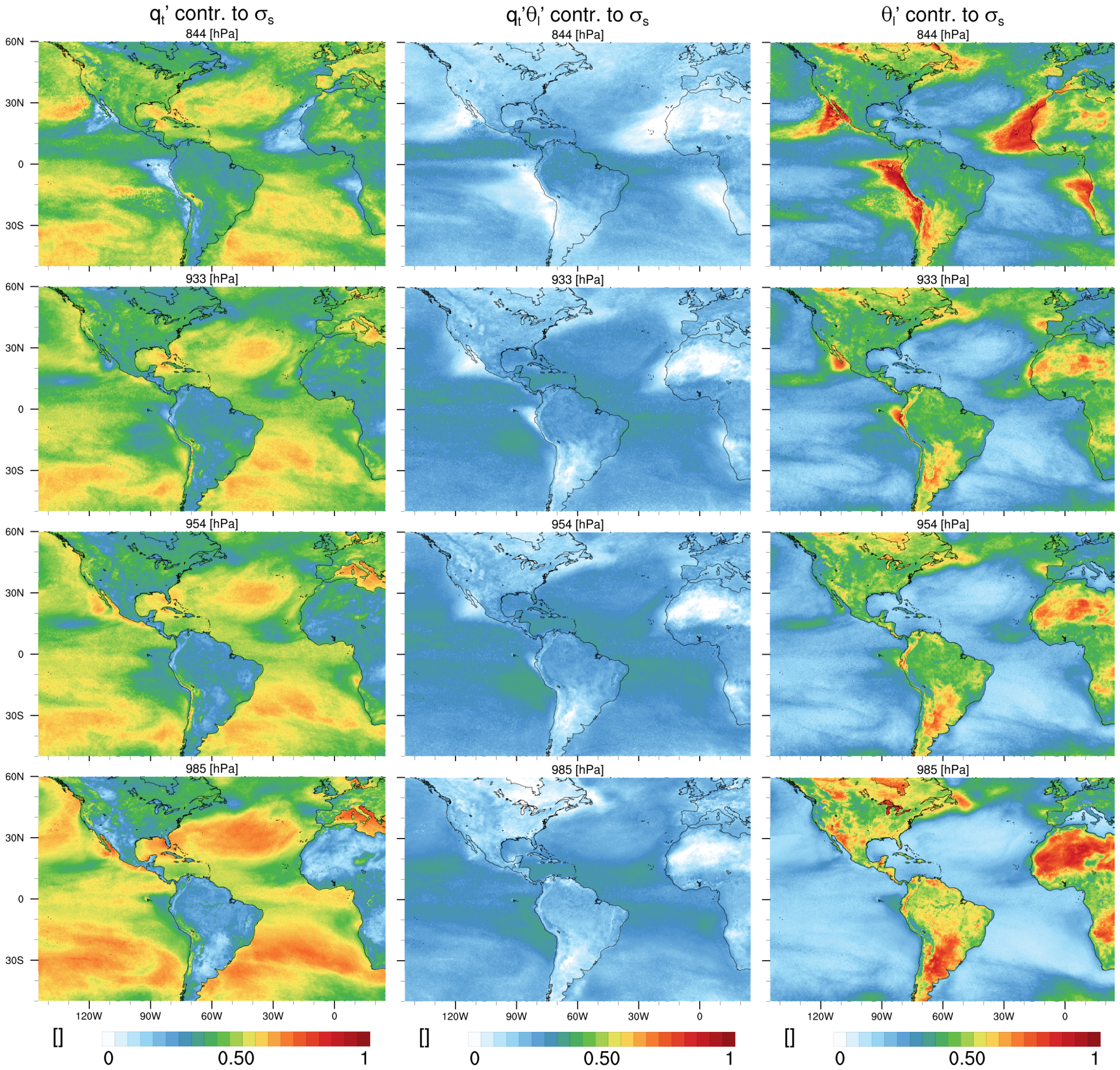


Figure 4.1.: Relative contributions of humidity q_t' (left), $\overline{q_t'\theta_l'}$ (middle) and temperature θ_l' (right) to the standard deviation of saturation deficit σ_s on subgrid-scale. Derived from instantaneous model output of a GFDL-AM3 CLUBB AMIP simulation for 30 day in May starting from NCEP nudged base run.

variability contribution is independent of the direction of heat fluxes. E.g. the Great Lakes in North America are colder than the surrounding land and the heat flux is directed into the water, independent of the diurnal cycle. Over land regions with high surface moisture, temperature variability has still the largest contribution, but in particular large river systems like the Amazonas show a higher humidity contribution. Over oceans the vast water availability allows latent heat fluxes to become be much larger, which in turn diminishes sensible heat fluxes. The humidity variability contribution dominates in the subcloud layer and then gradually decreases with height when clouds start to form and

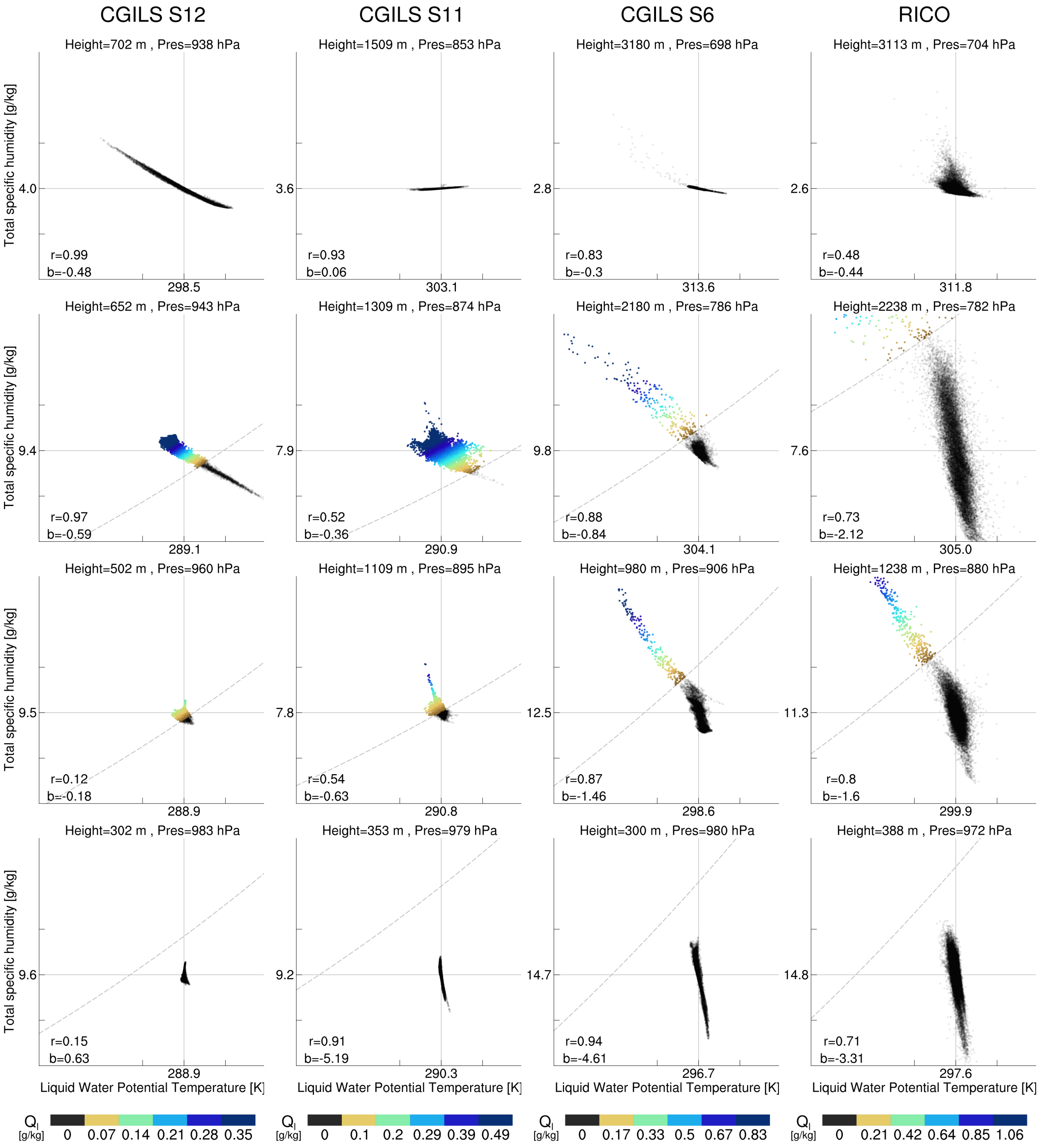


Figure 4.2.: Total specific humidity vs. liquid water potential temperature scatterplots of marine prototype LES cases. From left: S12 (column 1), S11 (column 2), S6 (col. 3) and RICO (col.4). Each row shows a similar vertical height level relative to the cloud layer: Altitudes increases starting from the bottom row. The color of dots indicate instantaneous the liquid water content. Grey dots are unsaturated LES grid-boxes. The dry saturation line in shown as dashed black line and the height level mean $q_t$and $\theta_l$is show as black horizontal and resp. vertical line with the corresponding value attached at the grid axis. The two other tickmarks have a spacing of ± 1 g/kg (q_t-axis) and 1 K ($\theta_l$axis). For each plot the Pearson correlation coefficient (r) and the slope of a linear fit (b [g/(kg K)]) is given.

condensation releases latent heat. In the marine subtropical stratocumulus subcloud and cloud layer the direct temperature variability contribution (term III) is marginal, while the indirect influence via its covariability with humidity (term II) plays an significant role.

At cloud top and above temperature variability contribution takes over. Moving to shallow cumulus along the trade-wind trajectory a stronger humidity variability contribution is seen at all heights while covariability remains important. Outside the subtropical belt synoptic events strongly control the contribution of humidity and temperature variability. Systematic differences become less obvious and covariability plays a minor role. In particular the extratropical cyclones in the storm tracks show approximately equal contributions to σ_s (Fig. 4.1).

Vertical structure of thermodynamic variability

In addition to the GCM simulation, results from LES simulations for prototype boundary layer cloud regimes in along the subtropical trade-wind trajectory are used: stratocumulus(Fig. 4.2, S12), transitional regimes (Fig. 4.2, S11), shallow cumulus (Fig. 4.2, S6) and deeper shallow cumulus (Fig. 4.2, RICO). The geographical reference to these cloud regimes is introduced in Fig. 1.3 and their conceptual phenomenology is outlined in Chapter 1 and further analyzed in Chapter 3. The use of LES simulations is beneficial because the parameterized subgrid-scale of a GCMs is explicitly resolved, which will facilitate the discussion of the underlying mechanisms. Still, the limited area and idealized forgings of the LES setup neglect synoptic influences. However, emerging meso-scale circulations for the RICO case (double periodic boundary conditions) partly account for the feedbacks between convection and their environment. Each q_t vs. θ_l scatterplot in Fig. 4.2 has the same axis-range to visualize the height- and regime dependences of their mutual fluctuations in form of coverage of the q_t-θ_l phase-space. Most striking the magnitude of σ_{q_t} and σ_{θ_l} increases from stratocumulus to shallow cumulus (from left to right in Fig. 4.2) and across the subcloud layer and cloud layer and strongly decreases above (from bottom to top).

The magnitude of the covariability $\overline{q_t'\theta_l'}$ can be formulated in terms of humidity variability σ_{q_t} and temperature variability σ_{θ_l}, scaled with their mutual correlation coefficient ρ_{q_t,θ_l}.

$$\overline{q_t'\theta_l'} = \sigma_{q_t}\ \sigma_{\theta_l}\ \rho_{q_t,\theta_l} \tag{4.1}$$

Eq. (4.1) in conjunction with the scatterplots (Fig. 4.2) provides a geometrical understanding how the thermodynamic covariability can alter σ_s. Assuming fixed σ_{q_t} and σ_{θ_l} a strong negative correlation coefficient (ρ_{q_t,θ_l}=-1) will increase σ_s because the q_t-θ_l-distribution will be aligned more perpendicular to the saturation humidity line (Fig. 4.2, grey dashed line). The other extreme, ρ_{q_t,θ_l}=1, will cause an alignment along the saturation line, thereby diminishing the contributions of σ_{q_t} and σ_{θ_l} to σ_s. Considering distinct height levels relative to the cloud layer, following mechanisms influencing σ_s can be noted from Fig. 4.2:

- Surface layer: Small-scale buoyant updrafts are driven by surface fluxes. Buoyancy results from warm air (sensible heat) and/or moist air (latent heat), both leading to a small positive correlation ρ_{q_t,θ_l}. Positive ρ_{q_t,θ_l} is limited to the lowest levels (not shown in Fig. 4.2), however in cases where cold and dry air is advected over warm water the vertical extent will be larger.

- Subcloud mixed layer: Across the subcloud layer the correlation turns negative due to A) the conversion of buoyancy from the warm and moist plumes into vertical ascent, which are cooled by expansion, and B) downward mixing of warm and dry air from the free troposphere. The remaining buoyancy of the plumes is primarily driven by higher water vapor (smaller density than dry air). The altitude where the correlation turns negative results from the balance of surface processes vs. downward mixing. E.g. in the absence of convective mixing, the stratocumulus case S12 exhibits positive ρ_{q_t,θ_l} correlations up to cloud base (seen from the evolution of correlation coefficient with height, not shown here).

- Cloud layer: At cloud base the cloud core temperature increases due to latent heat release. However θ_l is further reduced due to an overcompensation by the liquid water correction term (see Eq.(2.4)) and clouds remain virtually colder than their surrounding environment. Throughout the cloud layer $\overline{q_t'\theta_l'}$ remains negative.

  The shallow cumulus cases exhibit highly negative ρ_{q_t,θ_l}, which shows up as narrow bands of mutual fluctuations in their θ_l-q_t phase-space (Fig. 4.2). For a specific cumulus ensemble the axis has a distinct signature in terms of slope and scatter across the axis, which results from the cloud size dependent lateral mixing rates and buoyancy production via latent heat release. The axis of the un-saturated part is driven by more diffusive, i.e. turbulent, processes and is orientated along the zero buoyancy line. The enhanced decoupling of cloudy and non-cloudy portions in the upper cloud layer erodes the correlation coefficient.

  The covariability of the stratocumulus cloud layer is strongly influenced by the small spatial scales of cloud top processes, i.e. radiative cooling and local turbulent mixing. Gentle inhomogeneities in these processes create a higher scatter of covariability, which is seen in the S11 case where shallow cumulus clouds are penetrating into the stratocumulus (Fig. 4.2).

- Above clouds: Above the cloud layer the covariability remains negative and is primarily driven by meso-scale fluctuations and gravity waves. Furthermore remnants of overshooting convection show up in the vigorous shallow cumulus case (RICO).

4.3. Reduced σ_s effect on CF and q_l

The important result of the previous section, that subgrid-scale temperature variability generally increases σ_s in heights relevant for cloud formation, implies that neglecting the subgrid-scale temperature contribution will consistently reduce σ_s. This has important implications for the associated cloud properties, because CF and q_l are a function of s and σ_s (Eq. (2.18) and (2.19) in Chapter 2). How neglected temperature variability translates into CF and q_l biases is explored by comparing the original CLUBB simulations to a modified CLUBB version with no temperature variability.

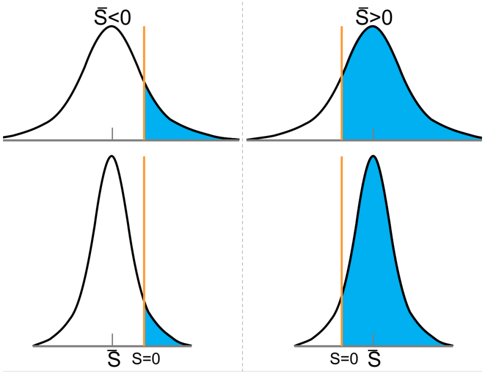


Figure 4.3.: Conceptual influence of reduced σ_s on CF (indicated by blue shading). Top row shows the s-PDF with a grid-box mean $\overline{s}$>0 (left) and $\overline{s}$<0 (right). Bottom row shows the corresponding PDFs with reduced σ_s, indicating the conditional influence on CF. The vertical orange line is the saturation humidity (s=0).

Fig. 4.5 shows the result of this comparison based on instantaneous model output of the GFDL-AM3 CLUBB simulations for 30 days. When σ_s is reduced, the spatial pattern of biases reveals a two-sided influence on CF, while for q_l a systematic reduction occurs. CF increases in regions that exhibit high CF already (e.g. stratocumulus regions), while CF decreases over regions where shallow cumulus prevail. CF differences are largest at heights where marine boundary layer clouds prevail. While the effects attenuates with height for CF, the q_l bias increases further. The two-sided influence on CF can be understood when considering the saturation deficit PDF, which is conceptually illustrated in Fig. 4.3. A reduction of σ_s is analog to a compression of its shape, which leads to a shift of non-cloudy portions of the PDF into the cloudy part (when CF>0.5 ↔ s>0), or a shift of cloudy portions into the non-cloudy part (for CF>0.5 ↔ s>0). Even though CF increases when σ_s decreases (for s>0), the grid-box mean q_l decreases, which results from the non-linear dependence of q_l on σ_s. When σ_s is reduced, the reduction of the second term in Eq. (2.19) dominates the increase of the first term (which scales with CF). This response is physically understandable: the compression of the s-PDF removes areas of high liquid water more effectively than compensated by the creation of new saturated areas with low liquid water. Fig. 4.4 illustrates the explicit scaling of CF and q_l in relation to σ_s, based on Eq. (2.18) and (2.19). Shown are typical examples of s and a slightly more extreme range of σ_s, than observed in Fig. C.11.

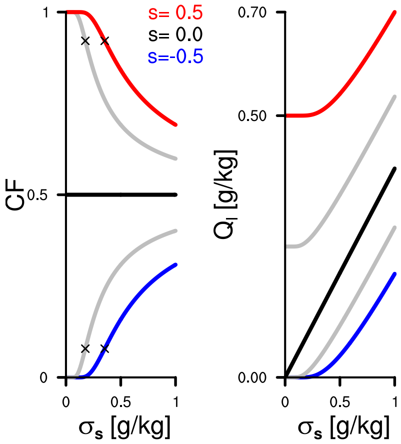


Figure 4.4.: Theoretical values of CF and q_l in relation to σ_s and for five different saturation deficit s (-0.5, 0.25, 0, 0.25, 0.5 g/kg) according to Eq.(2.18) and (2.19). Black cross: change in $\partial\text{CF}/\partial\sigma_s$ slope.

The scaling of CF and q_l in relation to σ_s can be summarized:

- $\partial\text{CF}/\partial\sigma_s$: CF increases monotonically with σ_s for s<0, respectively decreases for s>0 and has no dependence for s=0. $\partial\text{CF}/\partial\sigma_s$ is close to zero to small σ_s. Once σ_s exceeds a certain value, the dependence of CF on σ_s increases rapidly until a maximum is reached (indicated by the black cross on each the line in Fig. 4.4). For a further increase of σ_s, the slope inclines back towards zero. For $|s|$ closer to saturation the slope inclines to zero faster with increasing σ_s. Concluding, the sensitivity of CF on σ_s shifts to smaller σ_s for smaller $|s|$, which has important implications: The closer the grid-box mean is to saturation (s=0), the higher the sensitivity of CF on σ_s for small σ_s, and the lower for large σ_s.

- $\partial q_l/\partial\sigma_s$: q_l increase monotonically with σ_s, regardless of s. Similarly to $\partial\text{CF}/\partial\sigma_s$ the σ_s-dependence is close to zero for small σ_s and starts increasing once a certain value of σ_s is reached. The closer the grid-box mean is to saturation (s=0), the earlier the σ_s dependence starts. Different to CF, no maximum in $\partial q_l/\partial\sigma_s$ is attained, instead for increasing σ_s the relation asymptotically approaches the linear slope of $\simeq 1/\sqrt{2\pi}$, which results from the exponential contribution in Eq.(2.19) going to one.

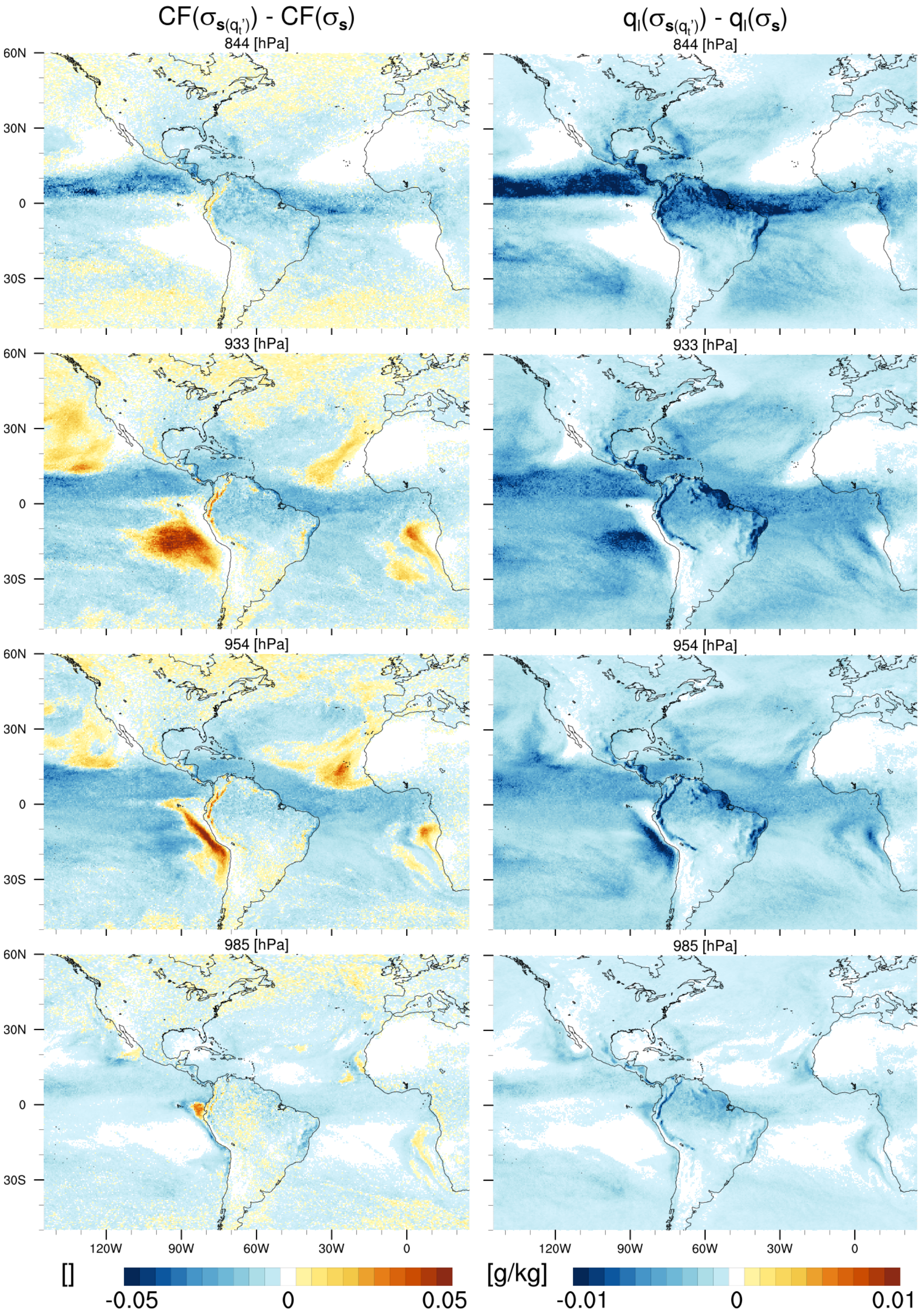


Figure 4.5.: Like Fig. 4.1, but showing the spatial pattern of the neglected subgrid-scale temperature variability σ_{θ_l}effect on cloud fraction CF (left) and liquid water content q_l(right).

4.3.1. Description of the s-σ_s-PDF

To further understand the relationship between σ_s and CF and q_l, in relation to s, their joint histograms are computed (Fig. 4.6), based on the instantaneous model output. The temporal frequency of occurrence reveals that: a) only certain combinations of s - σ_s occur. Saturation deficit is infrequently above 0.5 g/kg and has a long negative tail (truncated in Fig. 4.6), which indicates that grid-box supersaturation is effectively removed. b) There is a strong preference of occurrence for low σ_s (but not 0) values, along with slightly negative s. c) Grid-boxes with high subgrid-scale σ_s occur close to saturation, however their relative occurrence is low.

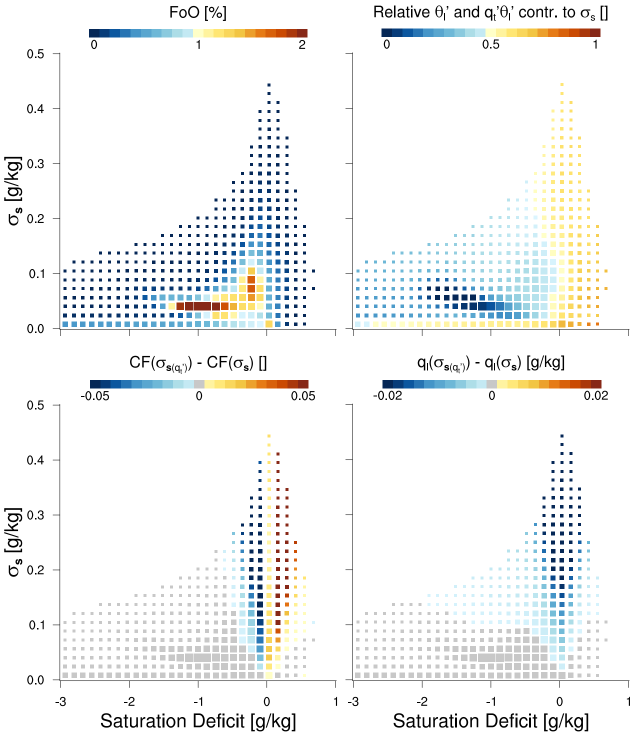


Figure 4.6.: Joint histograms of saturation deficit s (horizontal axis) and its parameterized subgrid-scale variability σ_s (vertical axis). Indicated in color: Top left: frequency of occurrence (percent of total) which is indicated as well in box size (for all plots). Top right: relative contribution of subgrid-scale temperature fluctuations to total saturation deficit variability. Bottom: bias in CF (left) and q_l (right) that results from neglecting subgrid-scale temperature fluctuations. Shown data: instantaneous model output of GFDL-AM3 CLUBB.

Considering the total contribution of θ_l' to σ_s (via covariability $\overline{q_t'\theta_l'}$ and directly), it shows that the higher the saturation deficit, the higher the contribution of θ_l'. Mapping the CF and q_l biases onto this joint histogram highlights where their largest biases occur, when subgrid-scale temperature is neglected. The CF bias is characterized by: slightly negative saturation deficit show highest negative CF bias, while slightly positive saturation deficit show highest positive CF bias. a) The strong dependence on s with indicates that slightly negative saturation deficit show the highest negative CF bias and vice versa for positive s. b) The positive relationship of σ_s with CF bias, i.e. the higher σ_s is the larger will be the CF bias. Relative to the dependence on s, the scaling of σ_s is weaker. c) The small envelop around s=0 where significant CF biases occur. With in increasing σ_s the s range is broadening. The q_l bias is in contrast not inverting with $s \pm 0$ and therefore consistently negative. Besides this major differece the scaling with s and σ_s is is similar to the CF bias with another exception that small but significant q_l biases can occur for more negative μ_s. One important conclusion is that the largest CF and q_l biases do not occur when the contribution of θ_l' to σ_s is highest.

4.4. Conclusion

Subgrid-scale temperature fluctuations influence the saturation humidity. The relevance of temperature variability for the cloud properties, i.e. cloud fraction (CF) and liquid water (q_l) can be studied using the saturation deficit framework, in which the temperature dependence on the saturation humidity is linearized around its mean state. Assuming a unimodal PDF, CF and q_l are function of saturation deficit mean (s) and variability (σ_s). Using this concept, the relevance of subgrid-scale temperature variability for cloud parametrization was split in two parts: first, the joint influence of subgrid-scale humidity and temperature variability on σ_s and second, the effect of neglected temperature variability on CF and q_l.

For this purpose GCM simulations using the advanced boundary layer transport and cloud parametrization CLUBB (Cloud Layers Unified by Binormals) are performed and idealized large-eddy simulation results are used to reveal the subgrid-scale fluctuations explicitly. Generally CLUBB in the GFDL-AM3 GCM shows promising spatial pattern of cloud properties, i.e. the stratocumulus regimes prevailing of the eastern subtropical ocean are well captured. Along the trade-wind trajectory, clouds transition towards more open convection together with ascending cloud tops as the marine boundary layer deepens. This behavior of marine boundary layer clouds is an improvement compared to results of the last CMIP5 (Guo et al., 2014).

Regarding the first question, saturation deficit variance σ_s is the linear combination of humidity fluctuations (q_t'), temperature fluctuations (θ_l') and their covariability ($\overline{q_t'\theta_l'}$) Global simulations show a strong land-marine contrast between the contribution of q_t' and θ_l' to σ_s. In marine regions the $\overline{q_t'}$ contribution dominates, while over land $\overline{\theta_l'}$ does. The contrast is strongest in the lower troposphere and scales with surface moisture availability, i.e. the associated surface fluxes. Higher up in the troposphere this contrast reduces, as cloud processes and circulation effects gain more importance. Following Tompkins (2003) temperature fluctuations are considered less important than humidity fluctuations, due to their shorter time-scales, as temperature perturbation can quickly be removed by buoyancy adjustment. However the subtropical marine cloud regime is in particular important because σ_s is large and temperature fluctuations have a significant contribution to σ_s. Especially in cloud relevant heights the impact of temperature variability is driven via its covariability $\overline{q_t'\theta_l'}$, while the direct temperature variability θ_l' contribution is small. The magnitude of σ_{q_t} and σ_{θ_l} increases along the trade-wind trajectory as convection get more vigorous and clouds transitions from stratocumulus to shallow cumulus. Throughout the cloud layer the covariability $\overline{q_t'\theta_l'}$ remains negative and highly correlated, which leads to a positive contribution to σ_s. From a process understanding point of view the positive contribution of negative $\overline{q_t'\theta_l'}$ to σ_s can be explained by the orientation of the q_t-θ_l-distribution relative to the saturation humidity line (Fig. 4.2). A negative correlation coefficient ρ_{q_t,θ_l}

increases the distance of q_t-θ_l-fluctuations from the saturation humidity axis, which amplifies σ_s. Therefore covariability $\overline{q_t'\theta_l'}$ has an important contribution to σ_s in the marine boundary cloud layer. Neglecting θ_l variability ($\rightarrow$ no covariability $\overline{q_t'\theta_l'}$ as well) will thus lead to significant reduction of σ_s.

Generally, the contribution of the covariability term to σ_s is positive, which has important implication for the cloud properties, addresed in the second question: A reduction of σ_s will further amplify cloud fraction CF differences that occur between stratocumulus and shallow cumulus regions, while systematically reducing grid-box mean liquid water content q_l. Reduced σ_s is analog to a compressed PDF, which shifts non-cloudy portions of the PDF into cloudy the part when CF>0.5 $\leftrightarrow$ s>0, and vice versa for CF>0.5 $\leftrightarrow$ s>0. The CF bias is largest in levels of warm marine boundary layer clouds. While the CF bias attenuates with height, the q_l bias further increases. Studying the formal relationship between s, σ_s and CF shows that the closer the grid-box mean is to saturation (s=0), the higher the sensitivity of CF on σ_s for small σ_s and the lower for large σ_s. The scaling of q_l depends as well on a threshold σ_s, which is lower for $|s|$ closer to saturation. One important conclusion is that the largest CF and q_l biases do not occur when the contribution of θ_l' to σ_s is largest.

Concluding, the results highlight the importance of temperature variability for cloud properties. The mutual relationships of temperature and humidity fluctuations indicate that their covariability plays a major role, which would facilitate an implicit treatment of temperature fluctuations in relation to humidity fluctuations (which is in line with results from Tompkins (2003) and Perraud et al. (2011)). A diagnostic treatment would reduce the complexity of the problem largely, but further studies, using synoptically driven cloud-resolving simulations would be beneficial to reveal the influence of cloud-circulation interactions in this context. In this context, the question what fraction of the temperature subgrid-scale variability gets resolved by the GCM grid-scales in relation to the GCM's horizontal resolution, is important and will be followed up.

Chapter 5

Large-scale evaluation

The motivation of this chapter is to explore how the subgrid-scale variability of cloud properties predicted by GCM statistical cloud parameterizations compares to large-scale satellite observations. This chapter build upon the theoretical concepts established in the previous Chapter 2 and the GFDL and ICON GCM simulations performed in Chapter 3.

Even tough the underlying fundamental physical concepts related to the Navier-Stokes equation are well established, a deterministic validation of the model's behavior is not feasible, due emergent atmospheric phenomena and complex interactions between parameterized processes. The evaluation of climate models is further complicated as there are no complementary observations available, i.e. no observational measurement technique can grasp the 3D atmospheric state at a particular point in time.

In the past, evaluation efforts mostly targeted either a) the mean global radiative balance of the GCM or b) small-scale processes in idealized conditions, with a few exception such as (Siebesma et al., 2004; Teixeira et al., 2011; Quaas, 2012; Dorrestijn et al., 2015). Traditionally one key motivation is to improve the mean global performance of GCMs. To improve the earth's radiative balance, parameterizations can be tuned, which has led to significant improvements. This approach is desirable for climate prediction efforts, carried out on regular basis such as the CMIP and IPCC (see Chapter 1). However tuning does not necessarily generate new understanding. Due to compensating errors in different parametrized processes, replacing one parameterizations in a tuned model is difficult as it initially often leads to degraded overall performance.

One prototype example is the cloud radiative effect. The pure knowledge about the existence of a model-observation-bias does not provide deeper insight on the origin of this difference, since the mean radiative properties are influenced by the overall cloud amount, the cloud macroscopic properties (mean LWP and its spatial arrangement) and cloud microphysical properties (effective radius) (Jakob, 2003). In order to support the development and implementation of physically more realistic parameterization the evaluation framework needs to give better guidance.

Another important aspect in context of error attribution is the difference between the overall model performance vs. the stand-alone behavior of the cloud parameterization. Comparing model simulations to observations on a climatological timescale includes various circulation feedbacks which tend to obfuscate the origin of cloud parametrization errors. Using the GCM in forecast mode might be a way to solve this problem. With increasing spatial resolution of GCMs they get more similar to numerical weather prediction (NWP) and eventually share technical infrastructure and model development (e.g. pioneered by the UK Met Office's unified model). The GFDL and the ICON model both benefit from such a development strategy, which is accompanied by better capabilities to perform realistic forecast simulations. Therefore in this work, short-term GCM simulations are performed and respectively the second day of the model integrations is used for a process based evaluation. Figure 5.1 shows the mean lower tropospheric stability ($\Theta_{700hPa} - \Theta_{1000hPa}$) (Klein and Hartmann, 1993) which is a well established metric for the boundary layer state. As expected for weather prediction models, but not self-evident

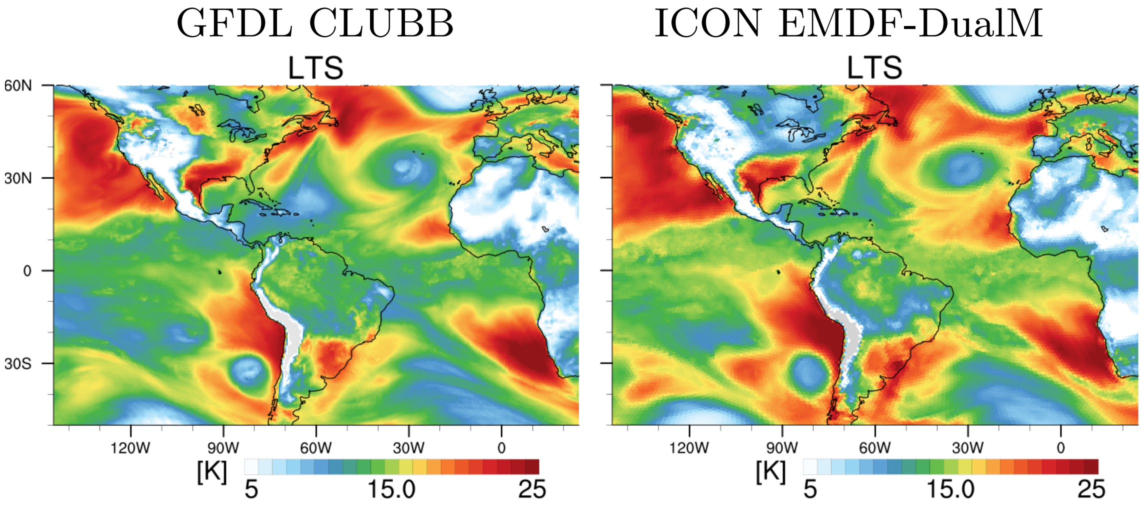


Figure 5.1.: Lower tropospheric stability for the GFDL-AM3 (left) and the ICON-GCM (right). Daily mean of the second day of a GCM-forecast simulation initialized with GFS (GFDL) and IFS (ICON) analysis data.

for GCMs, both models produce similar synoptic states, which facilitates the comparison of their respective cloud parameterization and allows a process-oriented comparison to the observational data, i.e. the relationships/state quantities that reflect the essence of the cloud parameterization can be revealed.

One alternative evaluation approach following the outlined credo, would be the use of single column models (SCM), in which one column of a GCM is simulated without the interaction of the neighboring columns. This method allows a precise analysis of the parameterization's behavior, however the difficult part is to provide realistic boundary conditions so that the parameterized clouds can be compared to observations. One promising approach, recently developed by Neggers et al. (2012), is to perform continuous SCM simulations alongside with LES simulations and tower observations. The continuous operational evaluation helps to improve the representativeness of process-level relationships on longer timescales, which are beneficial for a statistical evaluation of the GCM's subgrid-scale behavior.

Subgrid scale variability (regarding GCM grid-box scales) of cloud water is tightly connected with the non-linear properties of cloud radiative and microphysical properties (see Chapter 1). Previous studies linked the cloud homogeneity to cloud types (Pincus et al., 1999; Pincus and Klein, 2000) and analyzed the influence on cloud radiative properties (Barker et al., 1996; Barker and Wielicki, 1997; Oreopoulos and Davies, 1998) and on the autoconversion (Weber et al., 2011; Kawai and Teixeira, 2012; Boutle et al., 2014a). Oreopoulos and Cahalan (2005) used MODIS data to derive a global overview of cloud homogeneity.

In this context various metrics were established to address the spatial variability within clouds:

- Inhomogeneity parameter χ, based on cloud optical thickness τ (Cahalan et al., 1994) (overbar indicates the grid-box mean value):

$$\chi = \frac{e^{\overline{\ln \tau}}}{\bar{\tau}} \tag{5.1}$$

- Non-dimensional homogeneity parameter (NHP) using in-cloud μ_{LWP_c} (Barker et al., 1996; Wood and Hartmann, 2006):

$$\mathrm{NHP} = \left(\frac{\mu_{\mathrm{LWP}_c}}{\sigma_{\mathrm{LWP}_c}}\right)^2 \tag{5.2}$$

- In-cloud LWP dispersion, a metric promoted by this study:

$$\nu_{\mathrm{LWP}_c} = \frac{\sigma_{\mathrm{LWP}_c}}{\mu_{\mathrm{LWP}_c}} \tag{5.3}$$

These metrics of cloud homogeneity are generally comparable and can be converted into each other. The advantage of ν_{LWP_c} is that it remains defined for clouds that exhibit very little variability, while NHP goes to infinity when σ_{LWP_c} approaches zero.

The scope of this study is to connect observations of clouds spatial variability to the GCM cloud parameterization. Early ideas of such relationships were explored by Considine et al. (1997) by linking the LWP of marine boundary layer clouds to Gaussian probability density functions of cloud layer thickness. Using the principal concepts of statistical cloud schemes formulated by Sommeria and Deardorff (1977) and Mellor (1977), they were able to explain the observed relationships among the LWP-PDF moments. Wood and Hartmann (2006) refined this approach by assuming an adiabatic cloud layer and showed that utilizing a Gaussian PDF provides an explanation of the LWP-PDF relationships in stratocumulus clouds.

In Chapter 2 the relationship between the saturation deficit (s) and liquid water PDF were explored assuming a uniform, triangular and Gaussian PDF. It was further explored how the vertical arrangement of cloud layers might influence the statistics of the vertically integrated cloud field. Within a cloud layer unique relationships between cloud fraction and LWP-PDF moments were found, while the LWP-PDF of multiple layers is potentially more complex.

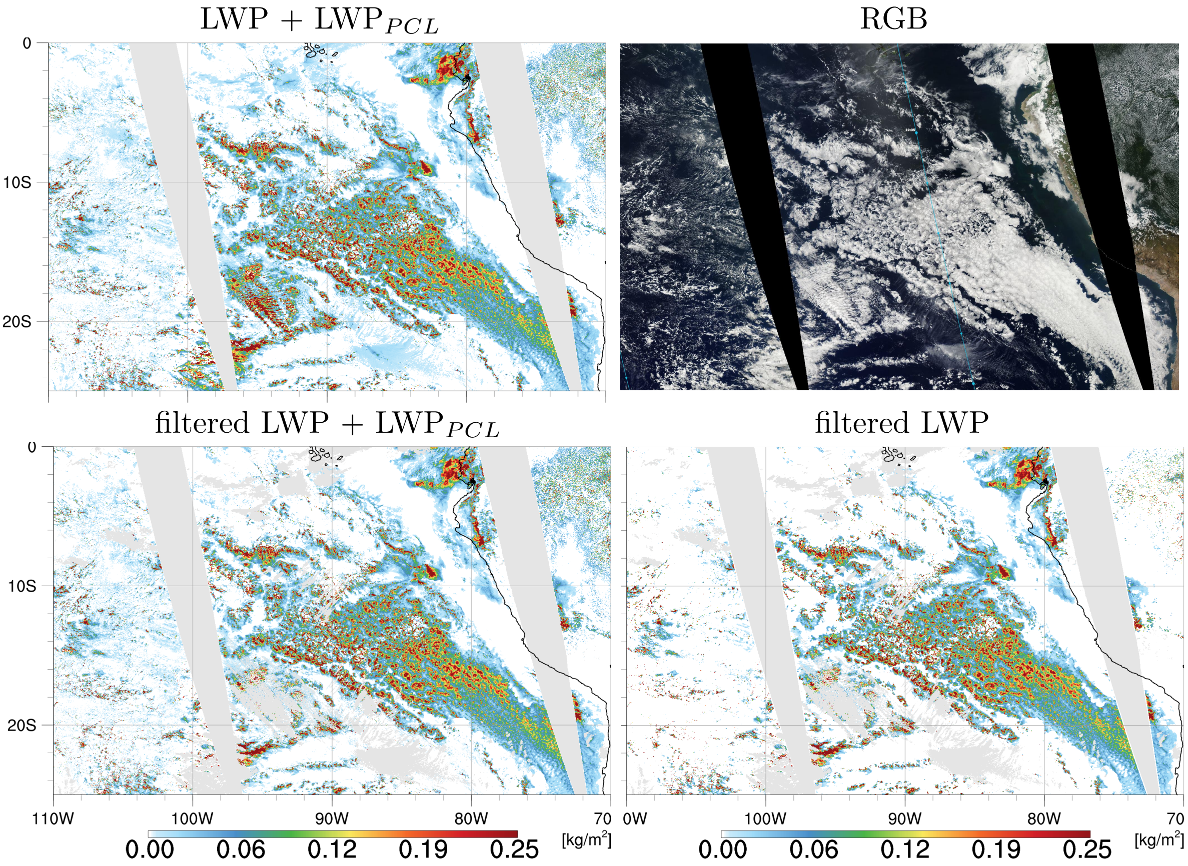
Using these relationships between the underlying s-PDF and its truncated saturated part, the evaluation of cloud parameterization can be facilitated by exploring the LWP-PDF moments Chapter 2. The LWP-PDF can be derived from satellites which has several advantages, essential for a GCM subgrid-scale evaluation, compared to aircraft or profiling measurements at one location. Low level orbiting satellites are able to capture vast domains almost simultaneously and their spatial resolution allows the detection of individual clouds.

The structure of this section is two-folded: First, methods to derive the LWP-PDF from observations are introduced. Second, the GFDL and ICON GCM forecast simulations are compared to the observed LWP-PDF. The comparison is of exploratory nature and focuses on the same day (25. Apr. 2013) and region (145°W to 25°E, 50°S to 60°N) as the GCM simulations discussed in Chapter 3.

5.1. Method

The GCM evaluation can be generally performed by either bringing the model on the same basis as the observations (evaluation in terms of observational parameters, i.e. radiances), or by deriving parameters from the observations that are closer to the model, i.e. cloud fraction. Based on the considerations of the previous section this work will facilitate observations that are closer to the subgrid-scale cloud properties predicted by the GCM cloud parameterizations.

The liquid cloud water path (LWP) PDF moments are derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) aboard the National Aeronautics and Space Administration (NASA) Aqua satellite on a sun-synchronous orbit in 705 km altitude (King et al., 1992). The MODIS level-2 satellite observations are used focusing on the



LWP + LWP_{PCL}

RGB

filtered LWP + LWP_{PCL}

filtered LWP

Figure 5.2.: Liquid water path (LWP) derived from MODIS collection 6, level-2 data with 1x1 km^2 nadir resolution and the corresponding RGB image with 250x250 m^2 resolution for the south-east Pacific stratocumulus region (25. Apr. 2013). The top-left figure shows the unfiltered LWP including the sub-pixel contribution. The bottom figures show the effect of the filtering which removes ice-phase clouds. The bottom right figure just contains the standard LWP from the level-2 cloud product, which misses the contribution from small cumuli over warmer ocean and over land.

daytime overpass which provide the essential visible channels.

In order to prepare the MODIS observations for a comparison to the GCM simulations several processing steps are required. To facilitate their discussion the steps are explained using a subtropical pacific sub-domain which contains all major prototype cloud regimes that will be discussed in this context (Figs. 5.2, 5.3).

MODIS collection 6 (C6), specifically the MOD06 cloud product, is used here (Menzel et al., 2015). The algorithms for the cloud optical properties, cloud top properties and cloud thermodynamic phase determination are comprehensively described in Platnick et al. (2003) and King et al. (2003) for collection 4 and with recent updates to collection 6 in Platnick et al. (2015).

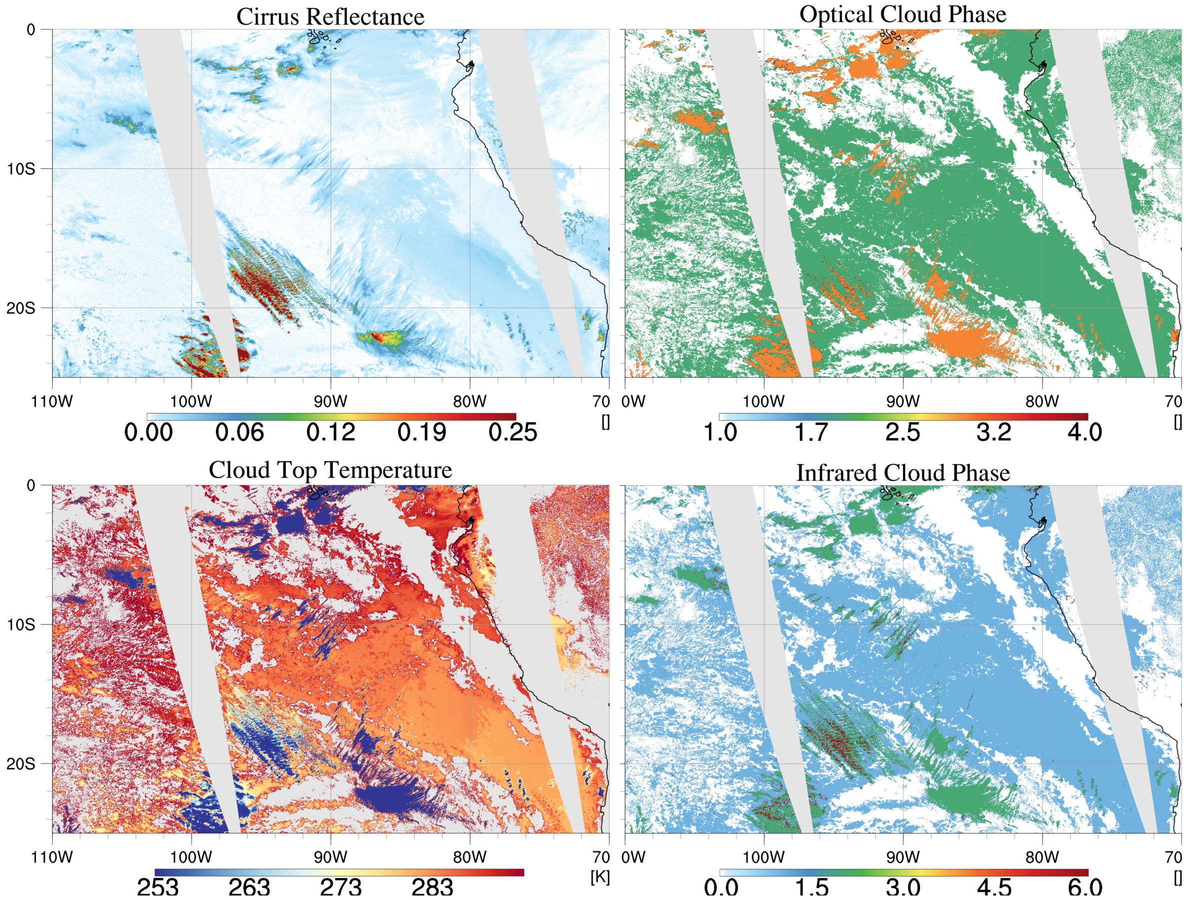


Figure 5.3.: MODIS level-2 cloud properties that are used in conjunction to filter out ice-phase clouds for the same scene as Fig. 5.2. Top-left: cirrus reflectance parameter. Top-right: optical cloud phase determination with the states clear (white), liquid cloud (green), ice cloud (orange), undetermined (red). Bottom-left: cloud top temperatures. Bottom-right: infrared cloud phase determination with the states clear (white), liquid cloud (blue), ice cloud (green), undetermined (red).

5.1.1. Liquid water retrieval from MODIS

The cloud water path for liquid and ice phase is derived using cloud optical thickness and effective radius, which are simultaneously obtained via Nakajima and King retrieval algorithm (bi-spectral solar reflectance method) described in Nakajima and King (1990) using multi-spectral reflectances.

The focus of this work is the evaluation of liquid boundary layer clouds, so the cloud water path is filtered for warm clouds in order to get an estimate for the liquid water path (LWP). Therefore data pixels potentially containing ice phase clouds are removed using the following filters, which are more rigorous then the original MODIS ‘subset to liquid cloud’ flag:

- Cirrus reflectance parameter using the 1.375 μm band. The scientific basis of the cirrus reflectance approach is described in Meyer and Platnick (2010). There is some cirrus reflectance basically over the entire domain, however the majority is sub-visible and influences the LWP retrieval not significantly. Applied filter: subset to pixel with values < 0.1 (dimensionless).

- Optical cloud phase determination using the 1.6 and 2.1 μm bands. Applied filter: use only liquid clouds, retain clear regions and discard undetermined cloud phase and ice clouds.

- Infrared cloud phase determination using the 8.5 and 11 μm bands. Applied filter: use only liquid clouds, retain clear regions and discard undetermined cloud phase and ice- and mixed-phase clouds

- Cloud top temperatures (CTT) using bands 11 μm. While CTT for clouds above 600 hPa is determined using the CO_2 slicing method is used, CTT for lower level clouds are determined directly via the IR window approach, which is new in MODIS-C6 and described in Menzel et al. (2016) section 3.1.1.b. Applied filter: subset to pixel $> 260K$.

The filtering variables are shown in Fig. 5.3 and reveal a coherent signature of cirrus clouds and their gross detection is achieved by either one of them. Still, looking at details shows that each variable is providing some additional value. Cirrus clouds can be identified by elongated patterns, deep convection shows up by its clumped signature and gradually deepening convection (cumulus congestus) gets visible by smoother horizontal gradients. In general, filtering is a balance between potential accuracy of the LWP retrieval and a loss in data. The filtering criteria described above yield a reasonable trade-off and sensitivity to the chosen thresholds is small.

5.1.2. LWP in broken cloud scenes

The exploration of the LWP retrievals over shallow cumulus regions with low cloudiness frequently shows missing pixels, even though other cloud properties (e.g. CCT) were successfully retrieved. This problem of partially cloudy scenes is well known for the clouds smaller than MODIS footprint of 1x1 km^{2}or in the vicinity of cloud edges (Koren et al., 2008).

Since MODIS-C6 the cloud product stores the cloud retrieval for partially cloudy scenes, i.e. the standard cloud water path does not include them but they are available from additional variable labeled ‘PCL’. The PCL variable contains the average result of the cloud retrieval over the 16 sub-pixels (250x250 m^2) that underly each 1x1 km^2 pixel. The use of the PCL variables requires some caution as some of the retrieval assumptions are violated such as the assumption of an overcast homogeneous cloudy field of view (FOV) Menzel et al. (2016). Therefore the quite restrictive choice of liquid cloud filtering criteria shows additional benefits. They ensure that scenes contaminated with cirrus or in the vicinity of deep convective clouds are removed, so that the remaining partially cloudy scenes are essentially shallow cumulus clouds. In this context, on the one hand, Zhao and Di Girolamo (2006) showed that using these sub-pixel information leads to an overestimation of cloud fraction for shallow cumulus regions, while on the other hand Seethala and Horváth (2010) showed that the cloud liquid water path of MODIS (not using the PCL information) is underestimated compared to the Advanced Microwave Scanning Radiometer-EOS (AMSR-E) for this cloud regime.

The scope of this work is the analysis of the spatial LWP variability over a 50x50 km^2 domain. The GCM grid-box equivalent LWP variability is derived over all cloudy and clear pixels. Therefore including the LWP from partially cloudy scenes is beneficial and will eventually lead to a smaller overall error than not including them, i.e. treating them as clear which is equivalent assuming zero LWP. Consequently the partially cloudy scenes are included in the in this work, following methodology of Brueck et al. (2015). Alternatively one could use an interpolation for pixels identified as cloudy by the MODIS cloud mask algorithm but removed from clear sky restoral algorithm (Platnick et al., 2003; King et al., 2003). This method, developed by Goren and Rosenfeld (2014) before the PCL information were in included in MODIS cloud product (prior MODIS-C6), uses the visible reflectance at 250 m resolution to fit optical thickness and LWP.

Further uncertainty of the MODIS LWP retrieval is introduced for the adiabatic cloud profile assumption, while in reality clouds, in particular open cumulus, are sub-adiabatic due to entrainment of environmental air into the cloud (King et al., 2003). However incorporating this effect into the LWP retrieval via a correction factor remains challenging (Merk et al., 2015).

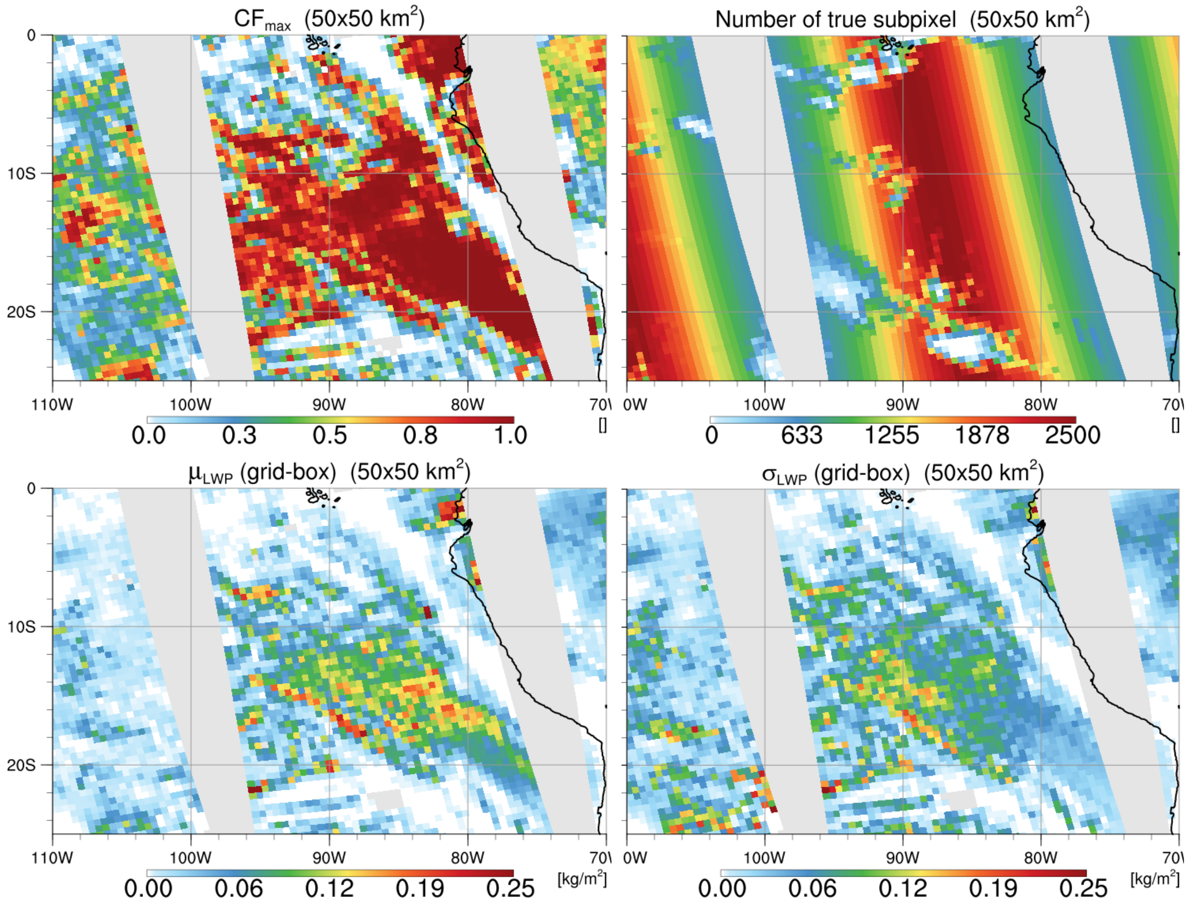


Figure 5.4.: Filtered MODIS level-2 LWP data including sub-pixel contribution which is sub-sampled onto domains of 50x50 km^2. At nadir each of these sub-sampling boxes contain 2500 pixels (same scene as Fig. 5.2. Top-left: Cloud cover (CF_{max}) derived from LWP > 0 pixels. Top-right: Number of valid pixel in each sub-sampling box. Bottom-left: Mean sub-sampled LWP resp. the whole sub-sampling box (μ_{LWP}). Bottom-right: Sub-sampled LWP standard deviation resp. the whole sub-sampling box (σ_{LWP}).

5.1.3. Data Handling

The introduced post-processing steps require the usage of MODIS level-2 data compared to MODIS level-3 data, which might be more convenient to use since level-3 data provides the histograms of LWP on a 1°x1° grid. In contrast level-2 data is stored in individual granules of five minutes overpass time on a bilinear grid. A regular granule with 1x1 km^2 resolution has the dimensions of 1354 x 2030 pixels (width x length) and a cross track FoV of 2330 km.

5.1.4. Sub-sampling

The moments of the LWP-PDF are derived over a spatial scale of 50x50 km^2, i.e. the full resolution MODIS level-2 data is aggregated or sub-sampled onto these boxes. At nadir each 50x50 km^2 contains 2500 pixels. With increasing viewing angle the level-2 resolution decreases, hence their number of pixels that fall in each 50x50 km^2 box (see top right of Fig. 5.4). The number of level-2 pixels per 50x50 km^2 box further reduces for regions where the quality filters were applied. In order to improve the quality of the sub-sampled dataset, areas containing too few pixels per 50x50 km^2 box are removed, specifically if there are less than 50 valid pixels or less than 25% of the possible pixels present. These quality criteria effectively limit the sensor zenith angle and reduce the width of the swath (see Fig. 5.3 and Fig. 5.4). This data loss is in regions of coarser level-2 resolution is further worthwhile, since clouds are increasingly observed from their sides rather than from their tops (Horváth et al., 2014).

Following the theoretical consideration of Sec. 2.2.6 the moments of the LWP-PDF are defined respectively to a) the whole 50x50 km^2 sub-sampling box (labeled grid-box) and b) the cloud area (labeled in-cloud).

5.2. Spatial distribution

The visual RGB image gives an overview of the multitude of cloud patterns. Overall the data loss due to high level cirrus and other ice phase clouds is low for the marine shallow cloud regimes and relativity high in deep convective regions associated to the Intertropical Convergence Zone (ITCZ) and storm systems (Fig. 5.5).

The sub-sampling procedure on MODIS level-2 is applied for the quasi-hemispheric domain (145°W to 25°E, 50°S to 60°N) to be congruent to the of the GCM simulations. Roughly 60 granules are stitched together along the swath to cover this domain. To fill the 2518 x 46 (along x across swath) sub-sampling boxes 123860 x 1354 level-2 pixels are

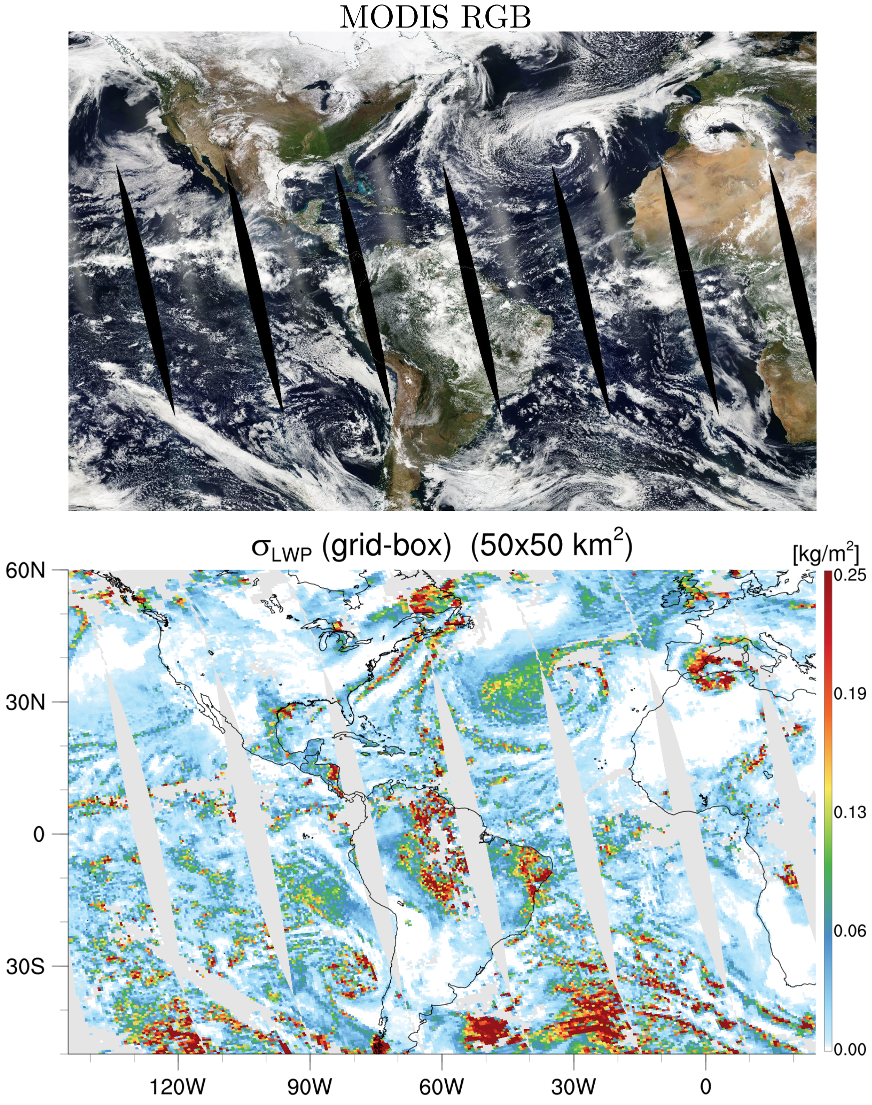


Figure 5.5.: Large-scale MODIS data of 60 stitched granules for 25. Apr. 2013. Top: RGB image with 250x250 m^2. Bottom: 50x50 km^2 sub-sampled LWP standard deviation resp. the whole sub-sampling box (σ_{LWP}) based on filtered MODIS level-2 data with 1x1 km^2 resolution.

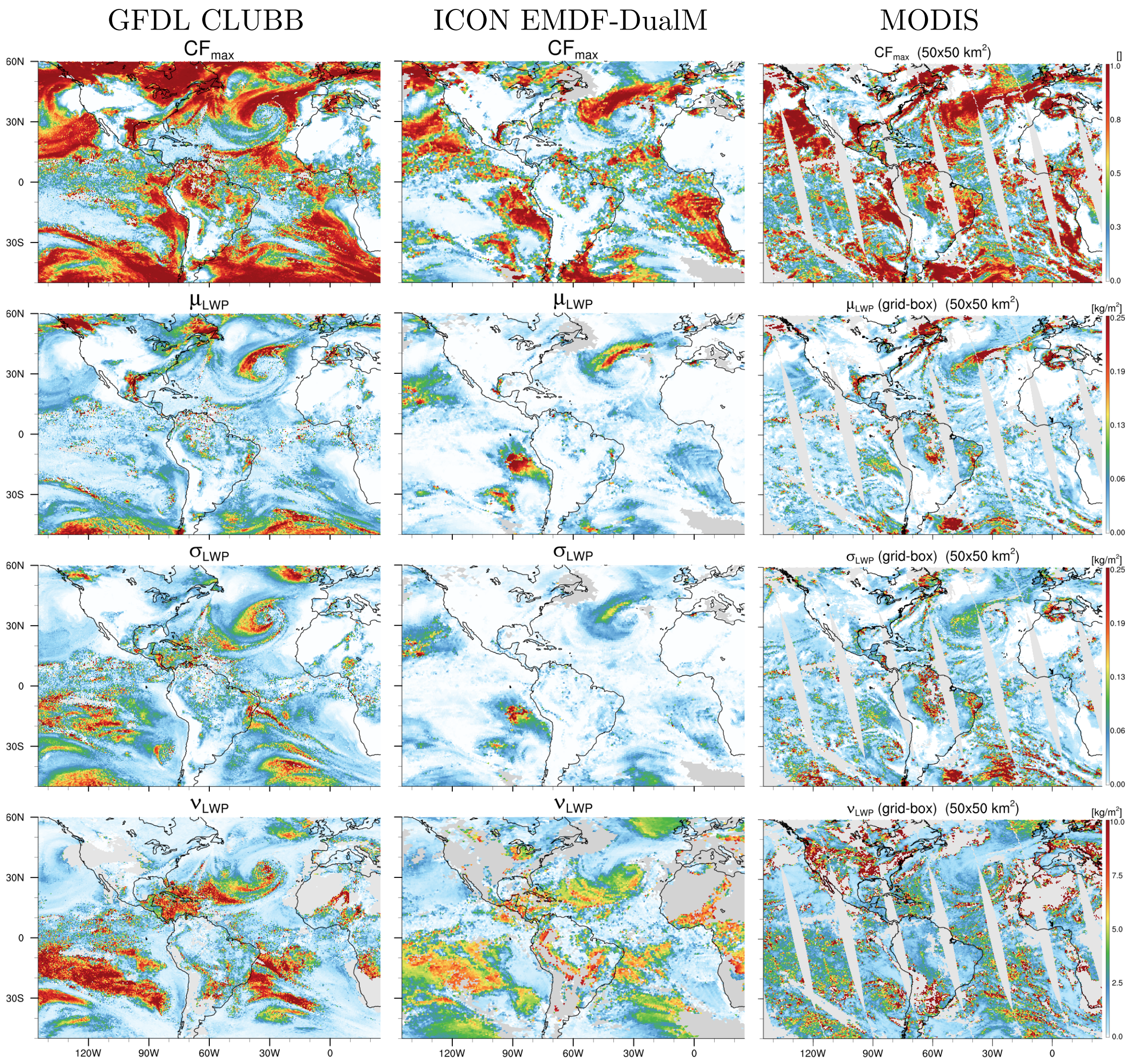


Figure 5.6.: Comparison of LWP-PDF statistics derived from the CLUBB cloud parameterization in the GFDL GCM (left), the EDMF-DualM cloud parameterization in the ICON GCM and the 50x50 km^2 sub-sampled filtered MODIS level-2 data for 25. Apr. 2013. The GCM show daily averages and MODIS the daytime overpass. LWP-PDF moments are respectively the grid-box, i.e. sub-sampling domain size, which includes the clear areas as well. Bottom row shows the dispersion ν_{LWP} which is the ratio of σ_{LWP}divided by μ_{LWP}.

used (see Fig. 5.5).

The sub-sampled statistics uniquely characterize different boundary layer cloud regimes as each different cloud type has a characteristic signature in terms of cloud cover, mean liquid water path and liquid water path variance. One important result, highlighted by Fig. 5.5, is the fact that the classical textbook cloud regimes do not occur. The stratocumulus decks exhibit complex internal structures with frequent clear patches. The intensity of the stratocumulus convection differs widely and is weakest for the Canary region in the North Atlantic (NEA) and higher off the Namibian coast in the South Atlantic (SEA) and off the coast of California in the North Pacific (NEP) and strongest off the Peruvian coast in the south-east Pacific (SEP). The more vigorous open cell SEP stratocumulus exhibit quite similar characteristics in terms of σ_{LWP} compared to the cold air outbreak in the North Atlantic (geographical references in Fig. 1.3).

Comparison to GCM

The GCM forecast simulations (24-48h) of the LWP-PDF statistics derived from the CLUBB cloud parameterization in the GFDL GCM and the EDMF-DualM cloud parameterization in the ICON GCM performed in Chapter 3, are now related to the MODIS subsampled LWP-PDF statistics. Generally, a one to one comparison is difficult because all 3 sources have their individual limitations, that lead to missing data in different locations for different reasons. For example the CLUBB scheme parameterizes the subgrid-scale properties of all clouds including deep convection and storm systems, while EDMF-DualM is an explicit boundary layer scheme and MODIS LWP observations are restricted to warm clouds. Therefore the results from CLUBB are subset to cloudy columns with a cloud thickness smaller than 3 km for this section, which removes deep convection. Comparing the position of major cloud features, i.e. fronts of mid latitude storms or cold air outbreak into the North Atlantic, between the MODIS RGB and the GCMs, the models perform reasonably well in forecasting the overall synoptic state. This is remarkable as they are just initialized by analysis data in contrast to the full data assimilation procedure that weather models have.

A general result, evident in the GCM simulations and the MODIS observation, is a general positive relationship between cloudiness, μ_{LWP} and σ_{LWP} (Fig. 5.6). The GCMs are able to capture the essential features of marine shallow convection which is primarily driven by closed cell stratocumulus decks in the eastern part of the Pacific and Atlantic basins, gradually transitions to open cell stratocumulus and further along the trade-wind trajectory to shallow cumulus convection. Specific conclusions can be made for the particular cloud regimes:

- Marine stratocumulus (Sc): the GCMs and MODIS agree relatively well in terms of spatial extent of the NEP Sc deck. Additionally they show spatial variability of cloudiness within the deck. Compared to MODIS, CLUBB produces a more homogeneous cloud deck while EDMF-DualM a more inhomogeneous one. The SEP Sc deck in CLUBB shows a transition from closed cell to open cell stratocumulus too early on the trade-wind trajectory (i.e. too south-eastward). The NEA Sc region features more broken Sc which is captured reasonably by EDMF-DualM, while CLUBB shows a larger closed cell cloud deck. The SEA exhibits the largest difference in this context. While EDMF-DualM produces a large Sc deck that starts right off the Namibian coast, MODIS just observes patches of connected Sc. In even larger contrast CLUBB produces Sc further away from the coast. Considering the LWP-PDF moments of the stratocumulus region EDMF-DualM overestimates μ_{LWP} and σ_{LWP} compared to MODIS, while the opposite behavior is found for CLUBB.

- Marine shallow cumulus: shallow cumulus covers vast regions of the warmer subtropical regions along the trade-wind trajectory. Zooming into the MODIS RGB visual reflectance (Fig. 5.5) reveals how the shallow cumulus fill the space between the organized convection towers. The sub-sampled liquid water path variance σ_{LWP} is fairly constant with values of around 0.05 kg/m^2. Large differences occur between CLUBB and EDMF-DualM. While CLUBB shows strong σ_{LWP} along the trade-wind trajectories compared to MODIS, EDMF-DualM misses or underestimates a large fraction of this cloud regime. MODIS frequently observes more organized and deeper shallow cumulus convection embedded in the trade-wind regime (zoomed detail in Fig. 1.3, middle), which are occasionally captured by EDMF-DualM. In contrast CLUBB produces deeper convective clusters that are spatially much larger (partly removed in Fig. 5.6)

- Marine cold air outbreak: the cold air outbreak in the North Atlantic advects cold arctic air over warm subtropical sea surface temperature (SST) (Fig. 1.3, region 5), which results in strong boundary layer convection. The center of rotation and the spatial extent are captured quite well by the GCMs. Compared to MODIS the associated distribution of μ_{LWP} and σ_{LWP} agree quite well (Fig. 5.6) (relative to the LWP statistics of the marine boundary layer cloud regime). Furthermore the models are able to resolve the fine-scale banded structure of the southward propagating convection.

- Shallow cumulus over land: unperturbed prototype small cumuli are predominant eastward of the Andes in the western Amazonas and in the region of Central Kalahari in South Africa. The LWP-PDF moments for vast cloud covered regions are well captured by CLUBB but strongly underrepresented in EDMF-DualM. Only in regions featuring more vigorous convection (central Amazonas) EDMF-DualM produces more realistic cloudiness but still shows too little μ_{LWP} and σ_{LWP}. Another region of relatively unperturbed shallow cumulus convection is present over Germany which will be discussed in detail in the following chapter in the context to the HD(CP)^2 project using cloud resolving modeling, see Chapter 6.

The last row of the Fig. 5.6 shows the grid-box liquid water path dispersion ν_{LWP} (ratio of σ_{LWP} divided by μ_{LWP}) which condenses the information of both. Stratocumulus regions exhibit low ν_{LWP} and the GCMs agree quite well MODIS. In shallow cumulus regions the

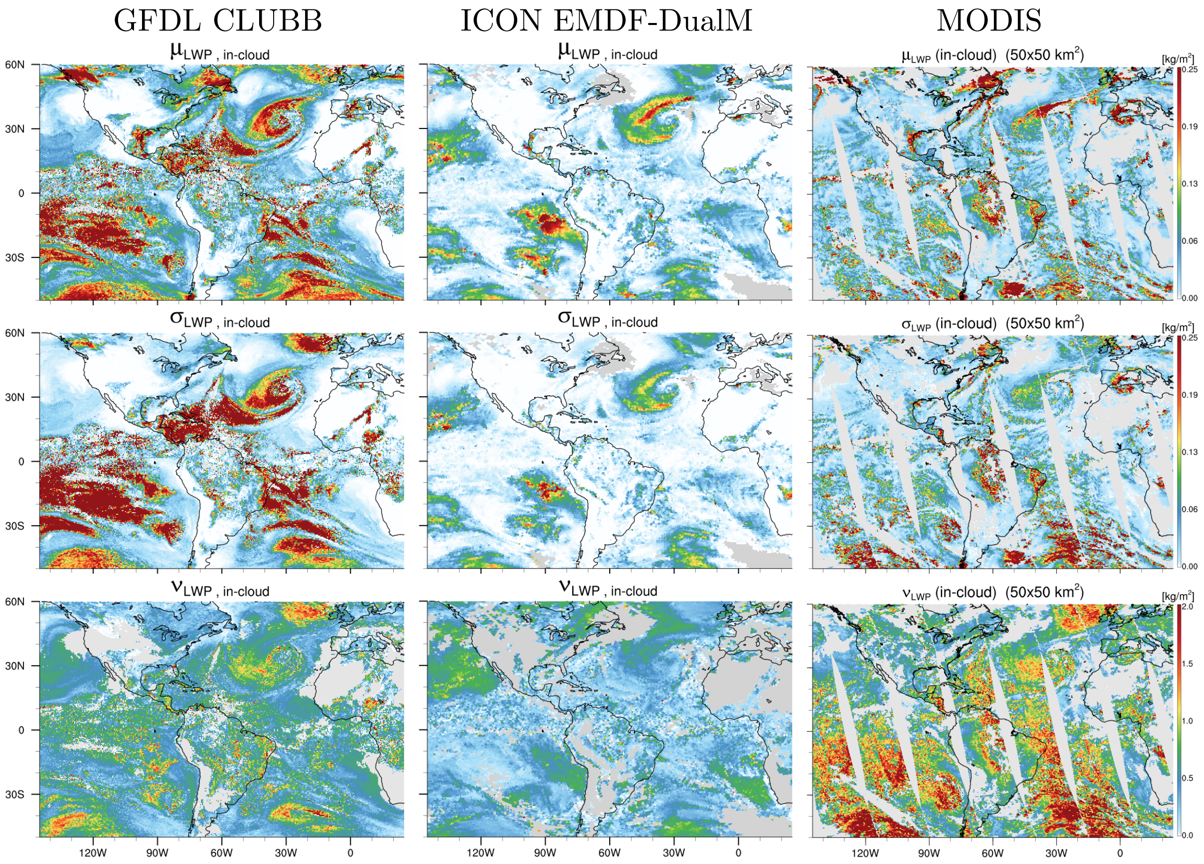


Figure 5.7.: Equivaltent to Fig. 5.6 but showing the in-cloud LWP-PDF moments, i.e. the LWP statistics are respectively the cloud area.

dispersion increases, i.e. σ_{LWP} increases stronger relative to the μ_{LWP}. CLUBB strongly overestimates ν_{LWP} compared too MODIS, while EDMF has slightly lower values but still overestimates. This relation might be influenced by MODIS not being able to detect optically thinnest small cumuli. Still, if the low LWP values would be missed, the ν_{LWP} would however increase since lower LWP values cause lower ν_{LWP}.

Differences that occur on grid-box scales are amplified when considering in-cloud values (Fig. 5.7). In particular CLUBB has a very small projected cloud area in marine shallow cumulus regions relative to the LWP which results in unrealistic high in-cloud values that are one order in magnitude higher than in MODIS. The in-cloud values of EDMF-DualM perform overall much better than on grid-box scale, which indicates a solid underlying conceptual idea but highlights the necessity for further tuning the overall cloud amount.

5.3. Conceptual relationships

In order to understand the relationships between the LWP-PDF moments and the cloud cover (CF_{max}) in more details their joint histograms are computed for the quasi-hemispheric domain (145°W to 25°E, 50°S to 60°N) (Figs. 5.8, D.12). While the geographical distribution of Figs. 5.6, 5.7 shows the daily mean value, for the following analysis instantaneous values of all time-steps are used.

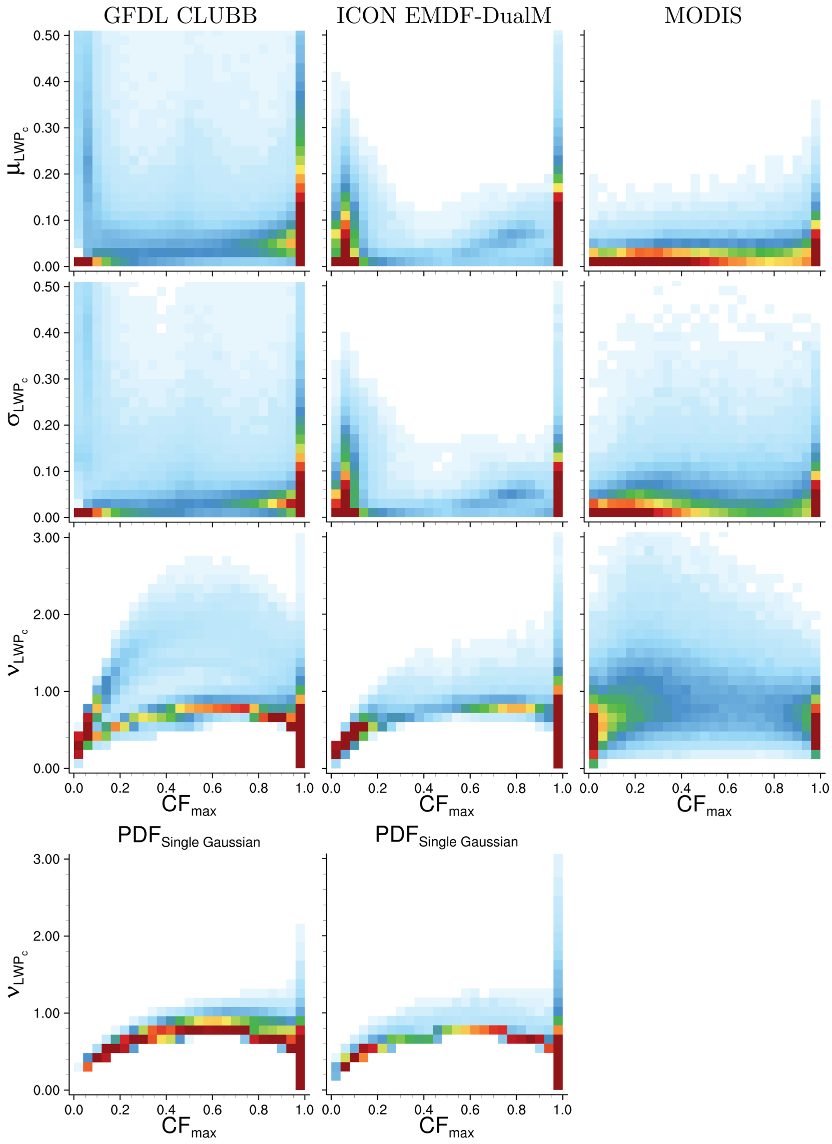


Figure 5.8.: Joint histograms showing the relationships between the in-cloud LWP-PDF moments and the cloud cover (CF_{max}). Instantaneous GCM output and all grid-boxes and MODIS sub-sampling domains for the large-scale domain (145°W to 25°E, 50°S to 60°N) for the 25. Apr. 2013 are used. Color show the frequency of occurrence (FoO): white 0%, blue ∼ 0.5%, green ∼ 1.0%, red ∼ 1.5%.

Compared to MODIS, CLUBB produces too high σ_{LWP} in particular for low CF_{max} which corresponds to the marine shallow cumulus cloud regime (Fig. 5.8). EDMF-DualM on the other hand overestimates σ_{LWP} at $\text{CF}_{max} = 0.1$ which results from the internally used threshold in the scheme. EDMF-DualM, and CLUBB to an smaller degree, produce either too low, or too high CF_{max} compared to MODIS.

Cloud cover is a powerful parameter for the cloud regime classification as it effectively determines the truncation fraction of the underlying distribution function (of total water, saturation deficit or cloud thickness like in Considine et al. (1997)). In particular when $\text{CF}_{max} = 1$ the underlying distribution is fully saturated (i.e. the shape is equivalent to the LWP-PDF). The μ_{LWP_c} can theoretically vary independently of σ_{LWP_c}, which results in a wide range of normalized shapes, which is summarized by their ratios, i.e. ν_{LWP_c}. Fig. 5.8 for $\text{CF}_{max} = 1$ shows this range of variability not only occurs theoretically. It is interesting that the original double Gaussian, as well as the diagnostic single Gaussian cloud parameterization produce a wider range of shape (ν_{LWP_c}) than that observed by MODIS.

Assuming a single layer cloud and a unimodal PDF (Chapter 2): If CF_{max} is even slightly less than 1, μ_{LWP_c} and σ_{LWP_c} get abruptly connected because the underlying PDF gets truncated, which implies that any further mutual variation of μ_{LWP_c} and σ_{LWP_c} goes only along with CF_{max}. Remarkably, this conceptual relationship remains valid for multi-layer clouds (see Chapter 2). The diagnostic single Gaussian still exhibits a close relationship of LWP-PDF moments with CF_{max}, which is slightly broadening as CF_{max} increases (bottom row of Fig. 5.8). Even though a) CLUBB and EDMF-DualM both use a double Gaussian and b) multiple cloud layers are overlying each other (which could potentially create all combination of μ_{LWP_c} and σ_{LWP_c} at any given CF_{max}), the majority of all columns is closely related to the single Gaussian. While EDMF-DualM just slightly exhibits more points with a higher dispersion, CLUBB features a second branch with higher dispersion that occurs with moderate FoO. This notable exception is introduced by deep convective columns in which the LWP variance is much higher relative to the mean, i.e. the LWP-PDF has a second mode in which the variance higher than what would be possible for a single Gaussian with the same mean. The MODIS observed relationships of ν_{LWP_c} with CF_{max} are generally much broader compared to a single Gaussian. Most striking is the difference at low CF_{max}, where the dispersion is not limited to narrow range of ν_{LWP_c}.

Skewness of LWP

Deriving LWP skewness (γ_{LWP}) and potentially even higher order moments from the sub-sampled MODIS level-2 data is straightforward. However with increasing order of the statistical moment, the more difficult and noisy gets its physical interpretation so the

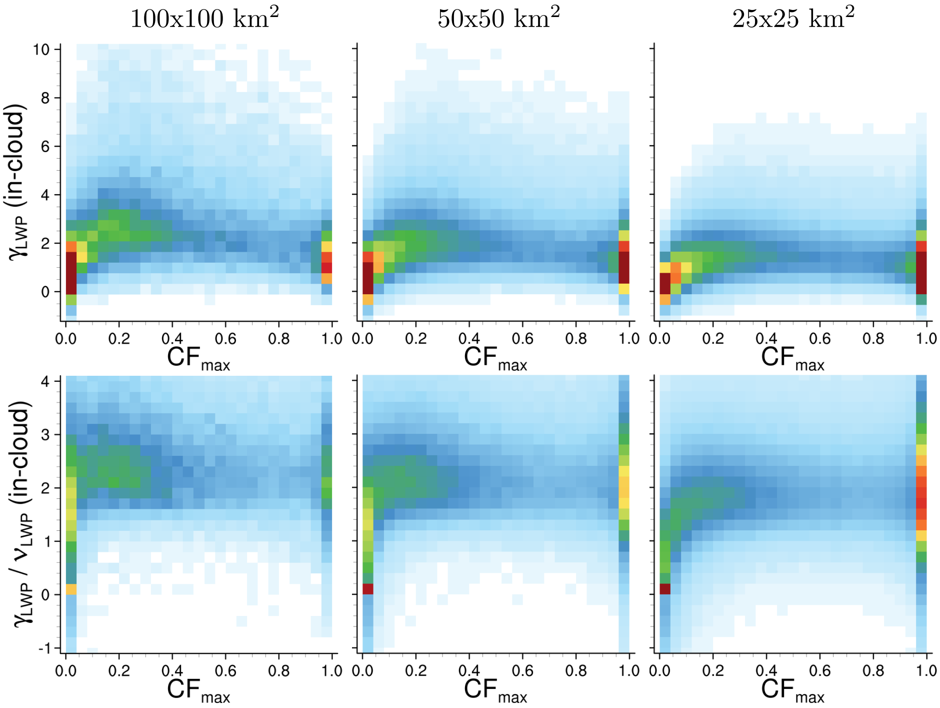


Figure 5.9.: CF$_{max}$-γ_{LWP_c}-relationship (top) and CF$_{max}$-$\gamma/\nu_{\text{LWP}_c}$-relationship for three different domain sizes derived from sub-sampled MODIS level-2 data (1 day, 145°W to 25°E, 50°S to 60°N). Color show the frequency of occurrence (FoO): white 0%, blue $\sim$ 0.5%, green $\sim$ 1.0%, red $\sim$ 1.5%.

discussion is limited to γ_{LWP} (Fig. 5.9).

For the 3 dimensional GCM data, the vertical integration of γ_{LWP} is analytically not possible and requires the use of a sub-column generator (Räisänen et al., 2004; Weber et al., 2011) which would be interesting but beyond the scope of this work. Still the analysis of liquid water statistics within one layer revealed a close relationship between dispersion and skewness (Chapter 2), which limits the potential of additional strong conclusions beyond the previous subsection. In this context the MODIS in-cloud ν_{LWP_c} is essentially similar to γ_{LWP_c} (Fig. 5.8 versus 5.9).

γ_{LWP_c} is small for small cloud cover. Considering the maximum FoO of CF$_{max}$-γ_{LWP_c}-relationship, one can infer a trend of increasing γ_{LWP_c} with CF$_{max}$ up to CF$_{max} \leq 0.4$. For higher CF$_{max}$ the slope becomes slightly negative, but the significance is smaller as the distribution of FoO shows a less coherent signal. Assuming a unimodal symmetric underlying PDF, like a uniform, triangular or single Gaussian (Chapter 2), CF$_{max} = 0.5$ marks the point where the associated LWP-PDF has a mode at 0 and transitions from a modal (CF$_{max} > 0.5$) to a non-modal (CF$_{max} < 0.5$). Hence for CF$_{max} < 0.5$ only the right tail of underlying PDF is saturated which explains the increase in positive skewness. Similarly to the relationship of CF$_{max}$ with μ_{LWP_c}, σ_{LWP_c} and ν_{LWP_c} for CF$_{max} = 1$ a wide range of γ_{LWP_c} occurs.

5.3.1. Effect of sub-sampling box size

When exploring the relationships among the higher order LWP-PDF moments one might question how the box sizes influence their relationships. In order to explore this potential aspect the MODIS level-2 data is additionally sub-sampled onto box sizes of 100x100 km^2 and 25x25 km^2 and LWP-PDF moments are derived (Fig. 5.9).

With decreasing sub-sampling box size, equivalent to a higher GCM spatial resolution, the main characteristics of CF_{max}-γ_{LWP_c}-relationship remain. The FoO of $\mathrm{CF}_{max} = 1$ increases as more cloud structures are completely captured by the smaller box size, but essentially the cloud field seems to be self-similar, following distinct power-law relationships (Neggers et al., 2003; Koren et al., 2008; Wood and Field, 2011; Rieck et al., 2014). Exploring the spatial variability of mesoscale cellular convection over marine boundary layer clouds, Wood and Hartmann (2006) conclude that these organized cloud structures on spatial scales of 10-50 km dominate the LWP variability. In this context it would then be interesting to further refine the sub-sampling box sizes to scales below mesoscale cloud clusters.

For large sub-sampling box sizes slightly higher γ_{LWP_c} occurs on average. The slope of the CF_{max}-γ_{LWP_c}-relationship in the range $0 < \mathrm{CF}_{max} < 0.4$ is reduced and shifted to lower skewness (Fig. 5.9). This trend goes along with a similar reduction in ν_{LWP_c}. The ratio of both, $\gamma/\nu_{\mathrm{LWP}_c}$, reveals some basic properties about the underlying PDF shape. While a triangular distribution is characterized by a constant ratio of 0.8 over a wide range of CF_{max}, $\gamma/\nu_{\mathrm{LWP}_c}$ of a Gaussian PDF monotonically decreases from 1.7 with increasing CF_{max}.

5.4. Conclusion

The evaluation of GCM cloud parameterizations is challenging because differences between model simulations and observations might result a) from a different atmospheric state in a model compared to the reality or b) from a deficient cloud parameterization itself. Furthermore an evaluation in terms of long-term statistical averages and spatial scales of the models horizontal resolution, just highlights regions where differences occur, but does not reveal the origin of the error.

In order to evaluate the subgrid-scale variability of cloud properties produced by GCM cloud parameterizations a novel approach has been explored which is characterized by a) short GCM forecast simulations which provide realistic synoptic situations and b) by comparing the models subgrid variability directly to equivalent observations, that resolve the model's subgrid-scale.

MODIS level-2 satellite observations with 1 km nadir horizontal resolution are aggregated onto boxes equivalent to the GCM grid-box size. Focusing on liquid boundary layer clouds the MODIS level-2 data was carefully post-processed and the LWP of partially cloud scene was included. Even though cloud retrievals in partly-cloud scenes have a higher uncertainty, they are able to capture the LWP contribution of small cumuli that cover vast regions over subtropical oceans and unperturbed daytime convection over land, which is supported by comparing the cloud field to the visible reflectances of MODIS.

Using MODIS satellite data has some distinct advantages compared to other observations: a) large domains are covered almost instantaneously, b) there is near-daily global coverage, c) relative to the covered area the spatial resolution is very high and the observations of small cumuli is possible, d) the MODIS cloud retrievals are scientifically well established and cross-evaluated against other observations, e) the MODIS instrument aboard the Aqua satellite orbits within the A-Train, which potentially allows to collocate spatial and vertical cloud information, f) a long data record is available so long-term or seasonal variations in subgrid-scale cloud statistics could be derived. Making use of these additional advantages that are beyond this work, will followed up in future research. However the disadvantages are that: a) the diurnal cycle cannot be studied (but using Aqua and Terra-MODIS observations for day and night (Gryspeerdt et al., 2014) or geostationary satellite data in conjunction would be one solution), b) the very small cumuli cloud cannot be observed.

In this work an exploratory evaluation focuses on one day and a near-global domain (145°W to 25°E, 50°S to 60°N) which covers the four major subtropical stratocumulus regions and their transitional areas to shallow cumulus further along the trade-wind trajectory into the ITCZ. Both, the GFDL-AM3 and the ICON GCM are able to produce a realistic synoptic situation.

In the first part of this section, a spatial comparison of the CLUBB cloud parameterization in the GFDL model and EDMF-DualM cloud parameterization in the ICON model are compared to the MODIS sub-sampled LWP statistics. Some striking differences in their bulk behavior are revealed: A generally positive relationship between cloudiness, μ_{LWP} and σ_{LWP} is found in the GCMs simulations and the MODIS observations. Both parameterizations produce extensive stratocumulus cloud decks. However, their spatial extent and LWP-PDF characteristics differ among each other and compared to MODIS. While EDMF-DualM overestimates cloud amount, μ_{LWP} and σ_{LWP}, CLUBB underestimates them compared to MODIS. The in-cloud differences are even more pronounced. Shallow cumulus convection is underestimated by both models, more severely by the EDMF-DualM over land. MODIS frequently observes more organized and deeper cumulus convection embedded in the trade-wind regime, that is just occasionally captured by the EDMF-DualM, while CLUBB produces deeper convective clusters with larger spatial scales than MODIS. These deeper convective cloud structures are challenging to parameterize as their spatial scales eventually reach sizes where the bulk mass flux assumption becomes invalid (convective area fraction becomes large relative to the grid size) and the convective grey-zone is entered (Dorrestijn et al., 2013; Boutle et al., 2014b). The example day in this evaluation features a cold air outbreak in the North Atlantic. The strong associated surface heating of cold air being advected over warm SST results in strong boundary layer convection which is well captured by both models. Compared to MODIS the associated distribution of the parameterized μ_{LWP} and σ_{LWP} agree well. Eventually the strong synoptic forcing leads to a more realistic simulated boundary layer compared to the less strong forced subtropical marine boundary layer cloud regimes, which is an interesting point for further investigation. Unperturbed prototype small cumuli over land are predominant eastward of the Andes in the western Amazonas and in the region of Central Kalahari in South Africa. The vast covered regions are well captured by CLUBB but strongly underrepresented by EDMF-DualM. Only in regions with more vigorous convection (central Amazonas) EDMF-DualM produces more realistic cloudiness but still shows too little μ_{LWP} and σ_{LWP}.

In the second part of this section, a more conceptual comparison of the GCMs and MODIS is done using the instantaneous GCM output. The LWP-PDF statistics are explored in terms of cloud cover. In general the conceptual relationships between LWP and cloud cover are quite well captured by the CLUBB and EDMF-DualM parameterization, however both schemes feature too extreme cloud cover, in particular low CF_{max} and close to 1 CF_{max} occur too frequently. Using these joint histograms the conceptual mechanisms of the cloud parameterization can be highlighted, in particular the cloud regime case switch underlying the EDMF-DualM parameterization (hard-coded at $\mathrm{CF}_{max} = 0.1$) produces unrealistic μ_{LWP} and σ_{LWP} statistics.

Cloud cover is a powerful parameter as it is highly connected to the shape of the LWP-

PDF underlying distribution function. This work goes beyond previous ideas of Considine et al. (1997) and Wood and Hartmann (2006), in which they used a Gaussian distributed cloud thickness model to explain the mutual relationships for stratocumulus. The results of this work are generally consistent and now extended towards a) the more complex double Gaussian PDF underlying CLUBB and EDMF-DualM and b) all boundary layer clouds. For the fully saturated case with $\mathrm{CF}_{max} = 1$, μ_{LWP_c} can theoretically vary independent of σ_{LWP_c}. When $\mathrm{CF}_{max} < 1$, the covariations are restricted by the truncation point of the underlying PDF. Even though CLUBB and EDMF-DualM both use a double Gaussian PDF which could theoretically result in broader spectrum of μ_{LWP_c}-σ_{LWP_c} combinations, they prefer shapes similar to a single Gaussian. In contrast the LWP-PDF statistics from MODIS show a broader range, that cannot be related to simple PDF-types in straightforward manner.

The scale dependence of the LWP-statistics on the sub-sampling domain sizes explored here (100x100 km^2 to 25x25 km^2) is weak. With decreasing domain size, equivalent to a higher GCM spatial resolution, LWP variance, dispersion and skewness are slightly reducing while their relationship to CF_{max} remain. This can be interpreted by a self-similar cloud field. Understanding how spatial cloud structures and convective organization are related on the process level to the LWP-PDF statistic, over varying domain sizes, would be very beneficial for the development of the next generation scale-aware cloud parameterizations (Schemann et al., 2013), which will be needed as the horizontal resolution of GCMs continues to increase.

Concluding, the observed LWP-PDF are more complex than what could be easily explained using simple unimodal and symmetric distribution functions. Still, considering the overall behavior of the CLUBB and EDMF-DualM cloud parameterization in comparison to the cloud schemes used during to fifth climate model intercomparision project (CMIP5), one can conclude that CLUBB and EDMF-DualM show a more physical realistic marine boundary layer structure. There is a vast potential for performance improvements of the CLUBB and EDMF-DualM which can be achieved by calibrating the different participating processes in the parameterization.

Chapter 6

Condensate variability in context to HD(CP)^2

This work is embedded in the ‘High Definition Clouds and Precipitation for Advancing Climate Prediction’ project (HD(CP)^2). The goals of this integrated project are: a) improving cloud and precipitation processes on a process level and b) developing and evaluating GCM cloud parameterizations. The motivation is largely driven by novel computational capabilities that allow cloud resolving simulation over domain sizes of $\mathcal{O}(1000\ km)$, which promotes internal cloud-circulation feedbacks. The progress should be achieved using the ICOsahedral Non-hydrostatic (ICON) atmospheric modeling framework (Zängl et al., 2015) (Chapter 3) that was extended towards large-eddy simulation model (ICON-LEM) applications (Dipankar et al., 2015) and the capability to perform synoptically forced simulations which produce a realistic atmospheric state (Heinze et al., 2016). These new modeling capabilities attach novel importance to high resolution observations, as they can now be directly compared and put in reference to the model results. Previously, differences in macroscopic cloud parameters, such as cloud cover or condensate amount, between models and observations were obfuscated in the subgrid-scale of cloud parameterizations. The resulting mutual stimulation of cloud resolving modeling and observations should in turn facilitate the understanding of GCM subgrid-scale variability.

The spatio-temporal variability of boundary layer characteristics and cloud properties produced by ICON-LEM is currently evaluated within the HD(CP)^2-community and presented in Heinze et al. (2016). This work contributing to this comprehensive evaluation effort by analyzing the spatial variability of vertically integrated cloud condensate. Furthermore the question in this section is, how the GCM parameterized cloud subgrid-scale variability relates to the explicitly resolved one.

6.1. Method

The ICON modeling framework was introduced in Chapter 3 in which global ICON-GCM simulations were performed. This section makes use of the ICON large-eddy simulations (ICON-LEM) performed in the HD(CP)^2 project (Dipankar et al., 2015). Currently four days in spring 2013 are simulated covering a Germany-wide domain. The modeling effort is ongoing so that more simulation days will follow. However the computational effort is so large that each simulation day needs approximately 1 month before the data becomes available to the HD(CP)^2-community. ICON-LEM uses open lateral boundary conditions and is nudged to the 'COnsortium for Small-scale MOdelling' (COSMO) numerical weather prediction model (Baldauf et al., 2011) with a horizontal resolution of 2.8 km. The horizontal resolution of ICON-LEM is refined by bisecting the edge length of the triangular grid in two steps from 625 m to 312 m to 156 m.

To put the ICON-LEM simulations in reference the intermediate-resolution COSMO model (Baldauf et al., 2011) is used again. The setup of these accompanying COSMO simulations (analog to Barthlott and Hoose (2015)) differs slightly to the ICON-LEM nudging COSMO data. Even though the horizontal resolution is coarser than ICON-LEM and a wide range of physical parameterizations are used, the COSMO model is a well evaluated and operational weather prediction model with high predictive skill.

Analog to the previous Chapter 3 the GCM forecast simulations with the GFDL-AM3 CLUBB- and the ICON-GCM EDMF-DualM cloud parameterization are used again for the 25 April. Similarly the MODIS level-2 cloud water path (CWP) and the subset to liquid phase only clouds (LWP) are used.

In order to compare the statistic of the LWP-PDF moments the approach of the previous Chapter 5 is followed, i.e. the MODIS Level-2 observations and the ICON-LEM- and COSMO simulations are sub-sampled onto 25x25 and 50x50 km^2 boxes. For the COSMO model data the explicitly resolved water on the 2.8 km grid and its subgrid-scale component are used in conjunction to derive the statistics on the sub-sampling boxes.

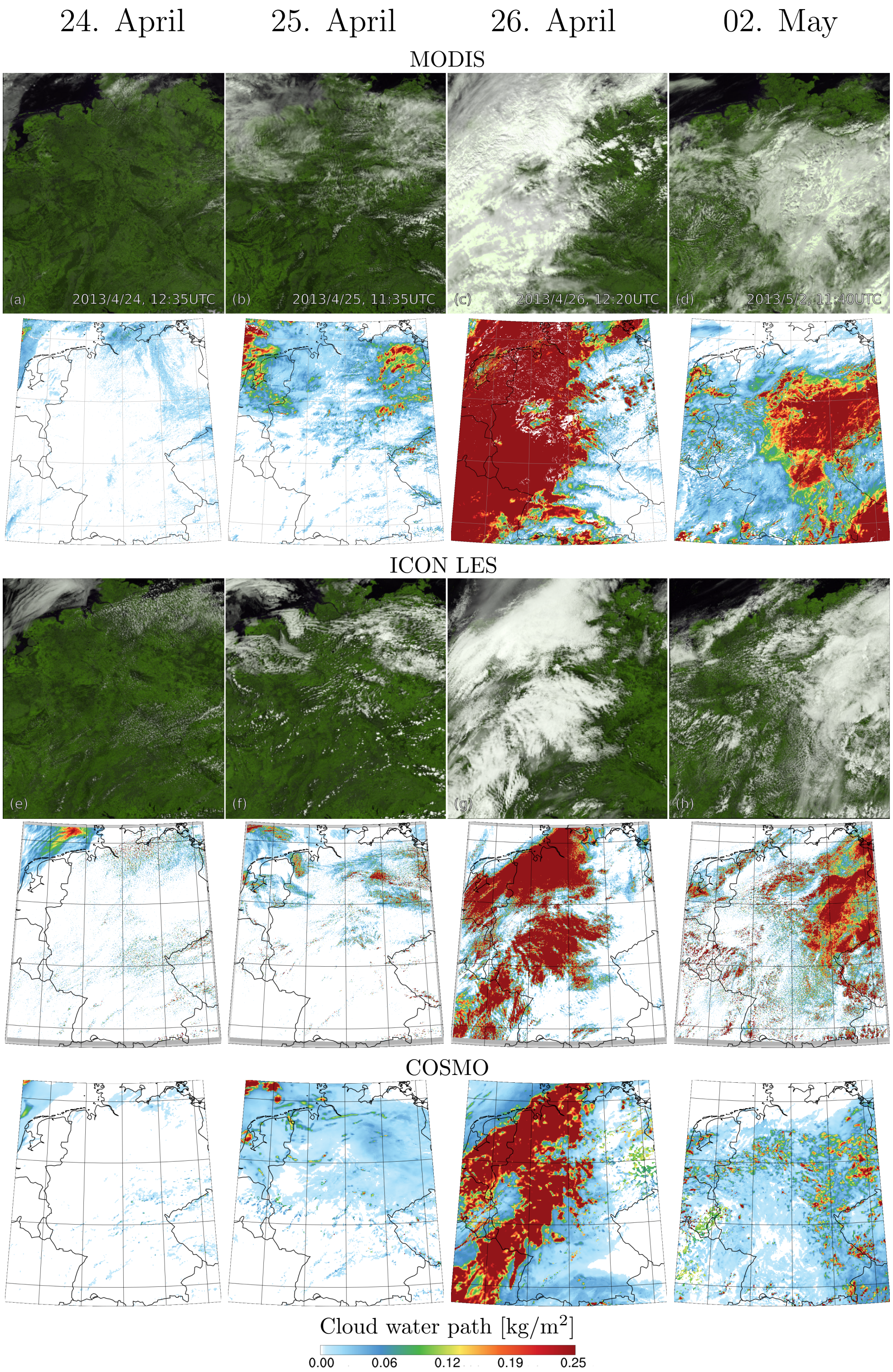


Figure 6.1.: Visible reflectance and cloud water path (CWP) of MODIS (row 1 and 2), synthetic visible reflectance and CWP of ICON-LEM (row 3 and 4) and CWP of COSMO. The four HD(CP)^2 simulation days are shown with times around noon when the Aqua-MODIS satellite overpass occurs. The visible images are produced within the HD(CP)^2 project by Leonhard Scheck and cited here from the HD(CP)^2 IOCN-LES evaluation publication (Heinze et al., 2016), i.e. a satellite forward operator is applied to the ICON-LEM 3D model output. Horizontal resolutions: MODIS: 250 m (row 1) and 1 km (row 2), ICON-LEM: 156 m or 625 m (row 3) and 156 m regridded to 1250 m (row 4), COSMO 2.8 km.

6.2. Overview of HD(CP)^2 high-resolution simulations

The HD(CP)^2 simulation days were chosen to cover a wide range of cloud regimes. Four ICON-LEM days are already available for post processing by the time of this analysis. Fig. 6.1 shows the MODIS, the ICON-LEM and the COSMO simulation results for the 24, 25, 26 April and 2 May 2013. The first two days (24 and 25 April) are under high pressure influence. On 24 April the lower troposphere is rather suppressed with small shallow cumuli in the north and east of Germany, while the 25 April is more convective throughout Germany and featuring isolated deeper convective events. On 26 April a front of a mid-latitude storm system is passing from west to east which is characterized by thick clouds and large scale organized convection. On 2 May the situation is similar to 25 April, but thick clouds are predominate in the eastern part of Germany.

Within the HD(CP)^2-project synthetic satellite images are produced using a satellite simulator (work done by Leonhard Scheck and cited here from Heinze et al. (2016)), which allow a direct comparison of ICON-LEM simulations to the 250 m reflectances observed by MODIS. The synthetic satellite images are produced from the ICON-LEM simulation on the finest grid (156 m), except for 25 April where the intermediate resolution output on the 625 m grid had to be used in consequence of output failure during runtime. Interesting to note from Fig. 6.1 is the representation of shallow cumuli. In particular on 24 April large differences between MODIS, ICON-LEM and COSMO are evident. While ICON-LEM is producing a homogeneous shallow cumuli coverage in the north-eastern part of Germany, MODIS has difficulties observing them as they can be barely seen in the visible MODIS image, while the cloud water path (CWP) including the partially cloudy 1x1 km^2 pixels detect them to some extent. Slightly different to the previous Chapter 5, Fig. 6.1 shows the CWP, which is motivated here by the compatibility to the accompanying analysis in Heinze et al. (2016). Nevertheless the spatial pattern of CWP and LWP are very similar since the ice-phase contribution is small (except for the frontal passage on 26 April), which is shown by the vertical profiles of ice-water content retrieved by CloudSat/DARDAR (see Fig. 18 in Heinze et al. (2016)).

Generally ICON-LEM is able to reproduce realistic essential features of the vertical and spatial distributions related to the boundary layer structures and of cloud parameters on these simulation days. While the skill of ICON-LEM in forecasting synoptic and meso-scale features is not superior to the well tuned COSMO model, the small and meso-scale variability of turbulence, water vapor and cloud water is much better resolved by ICON-LEM compared to observations, which predestines for GCM parameterization development (Heinze et al., 2016)).

While Heinze et al. (2016) targets the evaluation of all cloud related properties, the scope of this work is the spatial variability of integrated condensate. Analog to the previous chapter the CWP-PDF moments are derived by sub-sampling or aggregating the high-

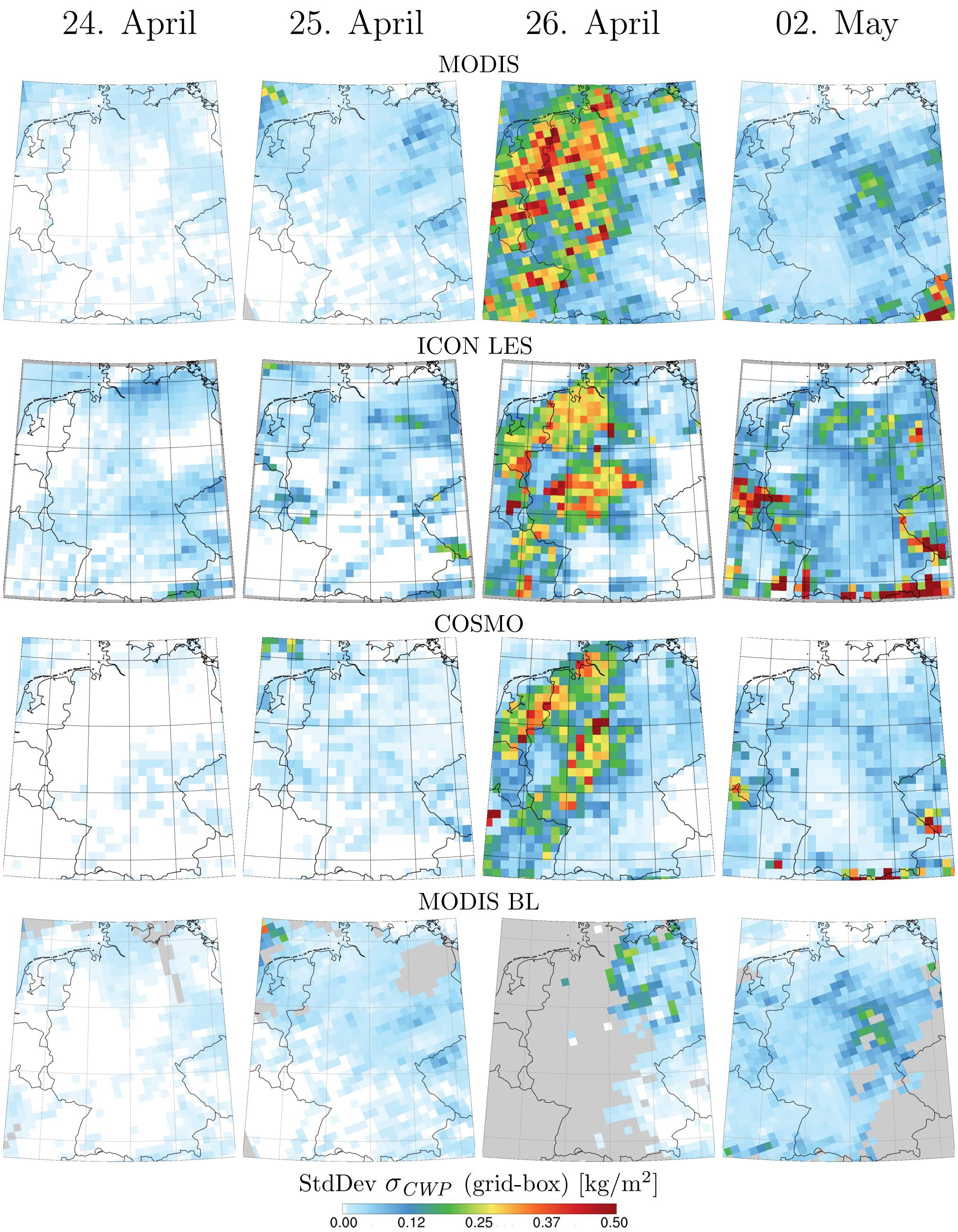


Figure 6.2.: Spatial distribution of cloud water path standard deviation (σ_{CWP}) over 25x25 km^2 sub-sampling boxes retrieved by MODIS with Δx: 1 km (row 1) and simulated by ICON-LEM with Δx: 156 m regridded to 1250 m (row 2) and by COSMO with Δx: 2.8 km (row 3). The bottom row (row 4) shows the σ_{LWP} (consistent to previous figures) using the filtered MODIS data where all possible ice-phase contamination is removed. This figure is incorporated in Heinze et al. (2016).

resolution data on larger box sizes. While the GCM evaluation used an aggregation onto 50x50 km^2 boxes (Chapter 5), MODIS, ICON-LEM and COSMO are now aggregated onto 25x25 km^2 boxes, which facilitates the spatial comparison and improves the attribution of variability to specific cloud regimes. Fig. 6.2 shows the spatial distribution of CWP standard deviation (σ_{CWP}) for the corresponding scenes of Fig. 6.1. Additionally the sub-sampling filtered MODIS data is included in the last row of Fig. 6.2 for which the ice-phase contamination has been removed analog to Chapter 5 (methods section), in order to provide consistency to the GCM comparison which will be performed in the following section. "It is observed that the signature carried in the 25x25 km^2 sub-scale variability of CWP is well represented by ICON with respect to MODIS. The simulations display a rich spatial pattern that can be attributed to the different cloud regimes that appear during each studied scene. By comparison, COSMO simulations have a smaller variability and exhibits little spatial gradients, but remain very consistent with MODIS retrievals" (Heinze et al., 2016).

Occasionally ICON-LEM σ_{CWP} exhibits higher variability in particular in the regions dominated by organized shallow convection, i.e. aligned convection in the southern half of Germany on 24 and 25 April. In that context the spatial variability is primarily driven by small cloudy areas of high CWP (CWP > 0.25 g/kg, beyond the color-scale), while MODIS and COSMO feature larger cloudy structures with less extreme CWP (conveys when zooming in Fig. 6.1). Notable is that the spatial variability of the filtered CWP from MODIS, i.e. σ_{LWP} is broadly consistent with σ_{CWP} in regions with little to moderate ice-phase contamination.

In this context an important aspect to the horizontal resolution dependence of the sub-sampling method, i.e. the grid-box size that is used to fill the sub-sampling boxes. In this context an important aspect to the influence of the horizontal resolution that is used

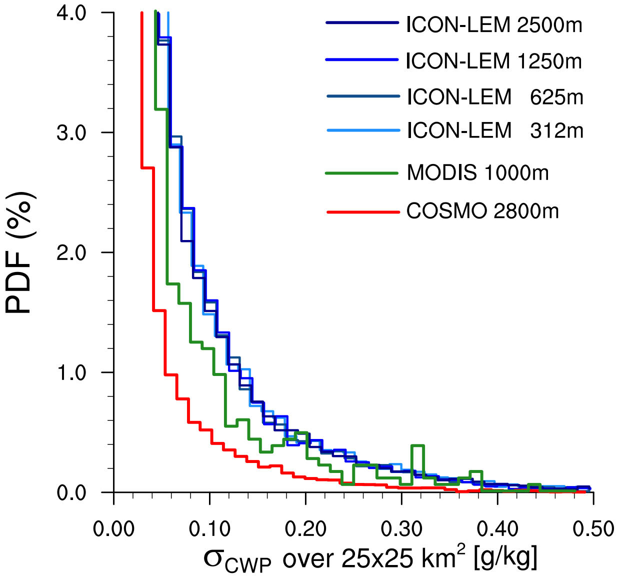


Figure 6.3.: Histogram of cloud water path standard deviation (σ_{CWP}) over 25x25 km^2 sub-sampling boxes retrieved by MODIS with Δx: 1 km (green) and simulated by ICON-LEM with Δx: 156 m regridded to 312, 625, 1250, 2500 m (blue) and by COSMO with Δx: 2.8 km (red) for 25 April.

to fill the sub-sampling boxes. In particular one might assume that the variability of ICON-LEM is higher than COSMO because of its finer resolution. However varying the horizontal regridding resolution of the original 156 m resolution of ICON-LEM, i.e. to 312, 625, 1250, 2500 m grids, does not significantly change the LWP variability on the scale of the sub-sampling boxes (25x25 or 50x50 km^2), which can be seen from the histogram of σ_{CWP} for 25 April (Fig. 6.1). However choosing a regridding resolution too coarse, i.e. in the limit the size of the sub-sampling box, will have an effect on σ_{CWP}. Furthermore the histogram reveals that the variability produced by ICON-LEM is consistently larger than COSMO and that σ_{CWP} of MODIS is between both models with a tendency towards ICON-LEM.

6.3. GCM simulation in context to HD(CP)^2

This section compares the LWP-PDF statistics from the CLUBB cloud parameterization in the GFDL GCM and the EDMF-DualM cloud parameterization in the ICON GCM with sub-sampled statistics of MODIS, ICON-LEM and COSMO. The comparison focuses on the 25 April 2013 since the synoptic situation on this day promotes widespread shallow cumulus convection, that can be compared best to the GCM cloud parameterizations (Fig. 6.1, column 2). Similar to the previous section the inter-comparison of the GCM cloud parameterization's behavior to high-resolution simulations is work in progress and will be extended to multiple days. In this regard the analysis has an exploratory character and is motivated by the question to identify auxiliary conditions of such an GCM evaluation.

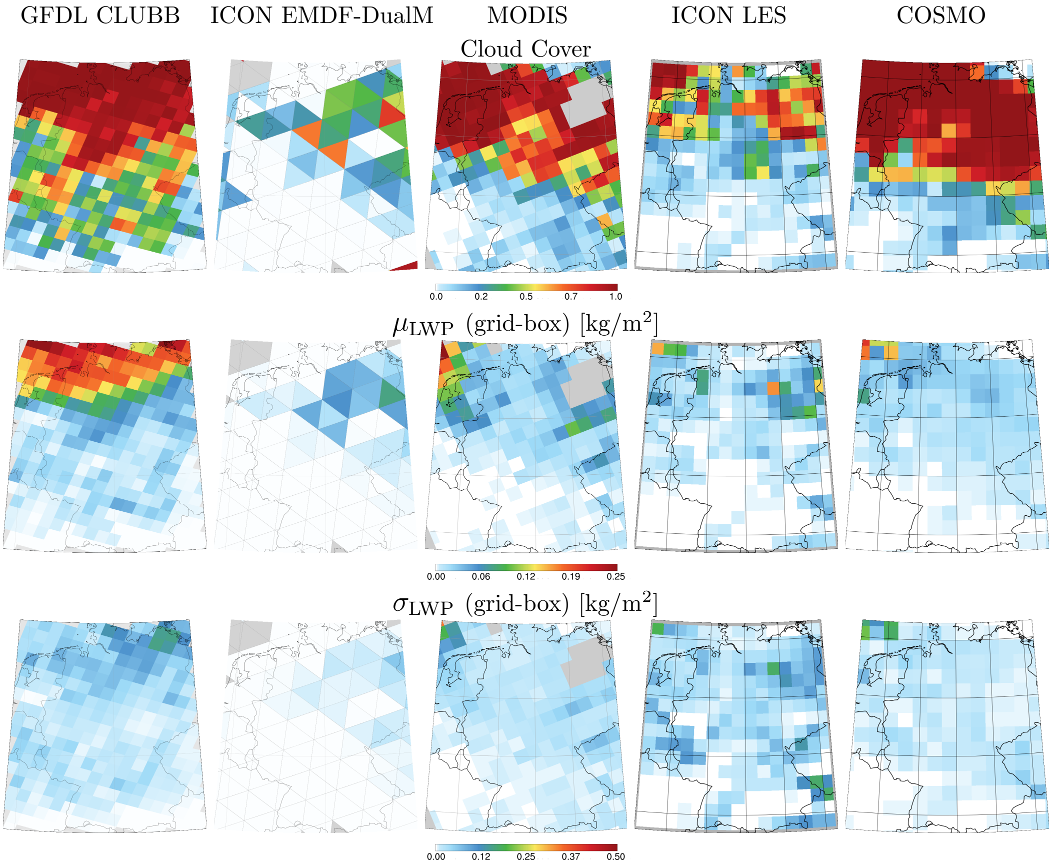


Figure 6.4.: Comparison of the GFDL-AM3 GCM using the CLUBB cloud parameterization (column 1), the ICON GCM using the EDMF-DualM cloud parameterization (column 2), sub-sampled MODIS level-2 data (column 3), sub-sampled ICON-LEM (column 4) and sub-sampled COSMO (column 5) for the 25 April 2013. The models show the daily average, and MODIS the daytime overpass. The horizontal resolutions are: 50 km for GFDL-AM3, 80 km for ICON-GCM, 1 km for MODIS, 156 m for ICON-LEM and 2.8 km for COSMO. The sub-sampling of the high-resolution data is onto 50x50 km^2 domains. Row 1: cloud cover. Row 2: liquid water path (LWP) mean μ_{LWP}. Row 3: LWP standard deviation σ_{LWP}.

In contrast to the previous section the MODIS data filtered for liquid-phase clouds, and the LWP of the ICON-LEM and COSMO simulation results are sub-sampled onto 50x50 km^2 boxes, in order to be consistent with the analysis of Chapter 5. Fig. 6.4 shows the spatial distribution of this comparison. Note that the left three columns are showing the results from the daily averaged GFDL-AM3 CLUBB and ICON EDMF-DualM GCM simulations and MODIS observations and is therefore a magnified version of Fig. 5.6.

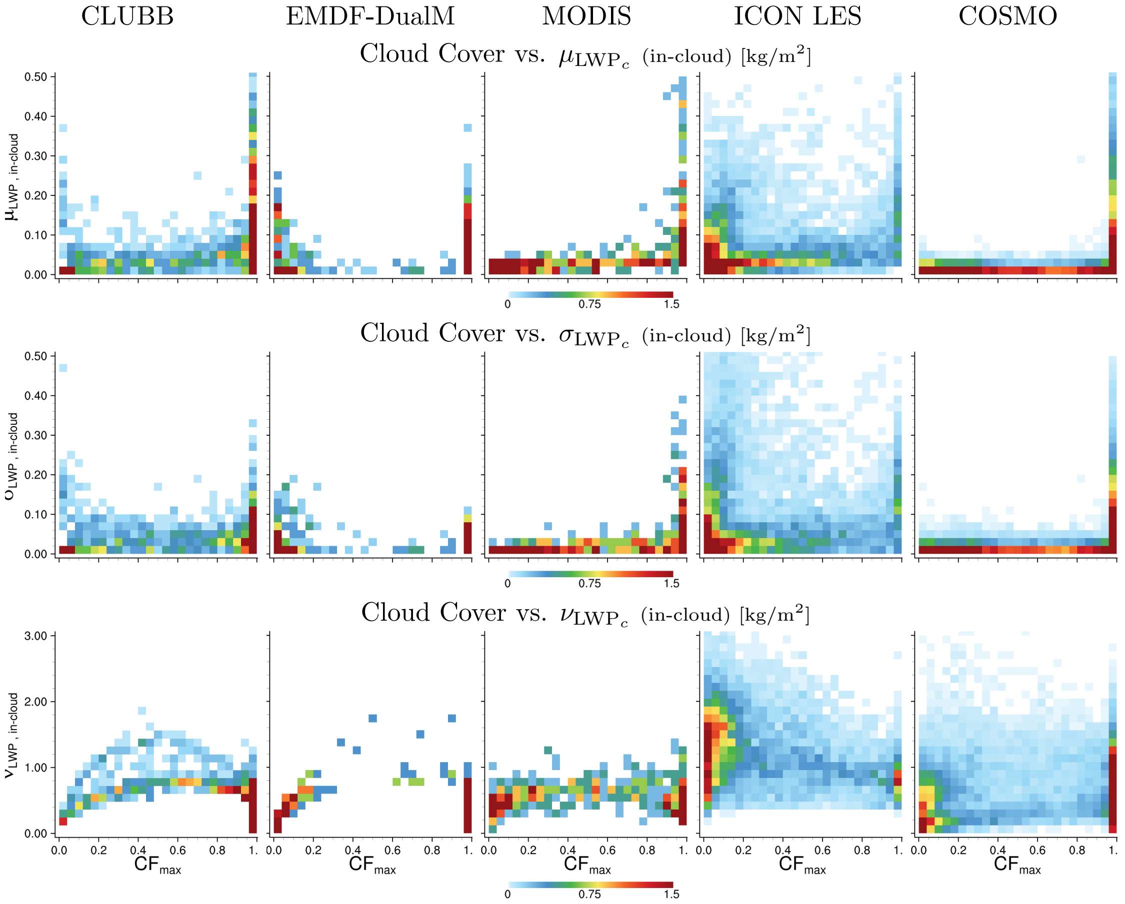


Figure 6.5.: Joint histograms of cloud cover (CF_{max}) vs. liquid water path LWP-PDF moments analog to Fig. 6.4 but using instantaneous data. Row 1: LWP-mean (μ_{LWP}). Row 2: LWP-standard deviation (σ_{LWP}). Row 3: LWP-dispersion (ν_{LWP}).

Cloud cover, mean liquid water path (μ_{LWP}) and variability (σ_{LWP}) produced by the GFDL-AM3 CLUBB simulation matches the general spatial distribution of the high-resolution ICON-LEM and COSMO simulation and the MODIS observation, even though cloud cover and μ_{LWP} are slightly overestimated. In contrast the ICON-GCM simulation with EDMF-DualM exhibits too little cloud amount and misses the shallow cumulus convection in the south-western part. Still, where significant cloud cover occurs the μ_{LWP} and σ_{LWP} are in the same order of magnitude compared to the other data sources. The deficiency of the LWP-PDF moments of ICON-GCM points out an essential problem that generally occurs for the evaluation of GCMs. The large-scale synoptic situation Fig. 5.5 and distribution of cloud properties Fig. 5.6 reveals that remnants of a frontal cloud system are present over the North Sea and northern Germany, which is insufficiently captured by the ICON-GCM. Therefore differences in the subgrid-scale distribution of LWP can likely be attributed to grid-scale deficiencies, rather than to the cloud parameterization itself. In particular shallow cumulus clouds are very sensitive to their driving heat fluxes and vertical thermodynamic gradients which further amplifies small differences in the synoptic situation.

Nevertheless analyzing the in-cloud LWP-PDF moments separates subgrid-scale behavior produced by the cloud parameterizations from the grid-scale deficiencies to some extent. Analog to the previous chapter the relationship between the in-cloud LWP-PDF moments and cloud cover are computed in order to reveal their mutual dependencies (Fig. 6.5). For the joint histograms of cloud cover and LWP-PDF moments the instantaneous values are used. The number of sub-sampling boxes therefore differs among the data sources. Since there is only one daytime satellite overpass for Aqua-MODIS the statistical population is smallest. ICON-LEM and COSMO provide every output 15 min while the GCMs just every three hours. One could potentially imporove the statistical significance by increasing the GCM output frequency or including the Terra-MODIS overpass.

The histograms in Fig. 6.5 summarize the parameterized and resolved LWP subgrid-scale variability which highlights several important conclusions: a) even tough the HD(CP)^2 domain is a small fraction of the quasi-hemispheric domain analyzed in the previous chapter (Figs. 5.8), GFDL-AM3 CLUBB and ICON-GCM EDMF-DualM show related LWP-PDF moment scaling with cloud cover. In particular σ_{LWP_c} is similar, which indicates that the relationship of μ_{LWP} and σ_{LWP} is characteristic for each cloud parameterization. b) The dispersion (ν_{LWP_c}) of MODIS is limited to smaller values compared to the large-scale domain (Figs. 5.8). c) The most striking difference occurs for the ICON-LEM in-cloud LWP-PDF moments at low cloud cover. μ_{LWP_c}, σ_{LWP_c} and ν_{LWP_c} are consistently larger than any other data source. For low cloud cover, large ν_{LWP_c} is accompanied with high γ_{LWP_c} which could be an indication that convective clouds remain too undiluted, i.e. they mix too little with their environment. d) The sub-sampled COSMO data exhibits little μ_{LWP_c} and σ_{LWP_c} for broken cloudiness, while for $\mathrm{CF}_{max} = 1$ higher values of both parameters occur more often. e) While the GCM cloud parameterizations favor a scaling similar to an assumed single Gaussian PDF, even though they are not forced to do so (see Fig. 5.8), ICON-LEM and COSMO do not exhibit such a distinct scaling. Their scaling is more complex and it is not possible to derive conclusion about their underlying distribution right away.

6.4. Conclusion

First results from the high-resolution ICON-LEM model in context to the HD(CP)^2 project are presented. Within the framework of HD(CP)^2 the ICON model has been extended to perform synoptically forced cloud resolving simulations which produce a realistic atmospheric state. While the ICON-LEM evaluation (Heinze et al., 2016) targets the general performance of boundary layer characteristics and cloud properties, this study focuses on the spatial variability of vertically integrated cloud condensate on GCM-scales explored via sub-sampling the ICON-LEM, the reference model COSMO and MODIS satellite observations. It is found that ICON-LEM is capable to match the satellite observations. Furthermore the different cloud regimes exhibit a distinct signature in terms of their spatial variability. In this context an important aspect are the different input resolutions for. In particular one might assume that variability of ICON-LEM is higher than COSMO because of its finer resolution. However the histogram σ_{CWP} reveals that a) regridding resolutions between 312 m to 2500 m of the original 156 m ICON-LEM does not significantly alter σ_{CWP} of the sub-sampling box and b) the variability produced by ICON-LEM is consistently larger than COSMO and that σ_{CWP} of MODIS is between both models with a tendency towards ICON-LEM.

Additionally the LWP-PDF statistics produced by the GFDL-AM3 CLUBB- and the ICON-GCM EDMF-DualM cloud parameterization are compared in the framework of HD(CP)^2 for the 25 April. Even though this day exhibits the least complex (i.e. most prototype) boundary layer moist convection case EDMF-DualM has deficiencies to capture the detailed synoptic situation. In contrast, CLUBB is closer to the high-resolution data. The conceptual relationship between the in-cloud LWP-PDF moments and cloud cover reveal a similar scaling of μ_{LWP}, σ_{LWP} and ν_{LWP} with cloud cover of the GCM parameterization compared to the quasi-hemispheric domain of the previous chapter. Particularly the ICON-LEM scaling is quite different, i.e. the in-cloud LWP-PDF moments are larger for low cloud cover. This behavior can be attributed to the smaller but more condensate loaded shallow cumulus clouds compared to MODIS and COSMO. This interesting characteristic is subject of current research and scientific discussion within the HD(CP)^2-community.

Currently ongoing, but beyond the scope of this work, is the extension of the GCM comparison for the other HD(CP)^2 simulation days. However it will remain challenging to attribute the subgrid-scale variability to the model's grid scale- or cloud parameterization's performance. Therefore potential alternative ideas are discussed in the general conclusion.

CHAPTER 7

CONCLUSION

This work is motivated by the question: how much complexity is appropriate for cloud parameterizations used in general circulation models (GCM). To approach this question, cloud parameterizations across the complexity range are explored using general circulation models and theoretical Monte-Carlo simulations. Their results are compared with high-resolution satellite observations and simulations that resolve the GCM subgrid-scale variability explicit. The key questions guiding the research within each chapter and their main results are summarized and subsequently concluded with respect to their to the initial question. than projected onto the initial question:

Chapter 2: How is the saturation deficit distribution connected to its cloud properties?

It is found that unimodal and symmetric PDFs exhibit an unambiguous scaling of liquid water mean and variability with saturation deficit variability, when binned according to cloud fraction. This characteristic is used to create a look-up stable to diagnose the s-PDF moments in terms of the q_l-PDF. A particular special property of unimodal and symmetric PDFs are their unique relationship between cloud fraction and liquid water dispersion (σ_{q_l}/μ_{q_l}) and skewness. For a given cloud fraction each PDF type produces one characteristic value, independently of the saturation deficit variability. Hence cloud fraction is a powerful parameter to determine the PDF's normalized shape.

In this line of motivation, analytical relationships for the column integrated cloud layer are developed and explored using Monte-Carlo simulations. The subgrid-scale liquid water variability within a vertical column is not additive (in contrast to the mean) because the spatial overlapping of condensate needs to be considered. Traditionally sub-column generators are used to calculate radiative and microphysical processes that depend on the vertically integrated liquid water variability. In this work the possibility of an analytical solution for the integrated liquid water variability was explored. The use of a covariance matrix, that contains all inter-layer covariances, provides a solution. Assuming maximum vertical overlap of cloud condensate, a methods to parameterize the inter-layer condensate correlations was established. Using the Monte-Carlo simulations, it was found that the inter-layer condensate correlations only depends on the cloud fraction difference between the layers. The results can be extended to multiple cloud layers and the use of a column bulk inter-layer condensate correlations promotes the application in GCMs.

These results of this section support the idea, that an evaluation of the cloud parameterization's subgrid-scale variability can be based on the cloudy part. Such an approach has intriguing capabilities, because on can take advantage of cloud resolving satellite observations.

Chapter 3: What are the characteristics of complex state-of-the-art cloud parameterizations in GCMs in contrast to simpler ones?

The advanced but relatively complex 'unified higher-order turbulent transport- and cloud parameterization' (CLUBB) (Golaz et al., 2002a; Larson and Golaz, 2005; Larson et al., 2012; Guo et al., 2015) and the 'eddy-diffusivity mass flux scheme' (EDMF-DualM) (Neggers et al., 2009; Neggers, 2009) are explored. Both schemes incorporate a statistical cloud parameterization which makes use of a double Gaussian PDF to predict cloud cover and liquid water. Even though the conceptual ideas are already existing for a few years, their performances have mainly been studied within single-column models in reference to idealized LES. Recently they were implement in GCMs, however their detailed subgrid-scale behavior of the underlying s-PDF and resulting liquid water statistics remained largely unknown. Short, synoptically driven forecast simulations were performed and produced similar atmospheric states which facilitates a process based evaluation. One key difference between both parameterizations is the treatment of shallow cumulus convection. In CLUBB they are represented by a partially saturated single Gaussian with large area fraction, while EDMF-DualM exhibits a fully saturated single Gaussian with a small area fraction. Even though both scenarios result in the similar cloud fraction, the associated cloud properties are different, in particular the in-cloud liquid water skewness show a different behavior.

A method to diagnose all related PDF parameters was development. Although both parameterizations are quite encapsulated in their host models, this was achieved in a consistent manner which provides the opportunity to reproduce the mutual relationships between the s-PDF and the cloud properties. Using the grid-box mean properties of the envelope s-PDF, simpler unimodal cloud closure were coupled to both models diagnostically.

The behavior of the simple diagnostic schemes (uniform, triangular, and single Gaussian), that were introduced in Chapter 2, is explored and compared relative to the double Gaussian PDFs of CLUBB and EDMF-DualM. Generally, differences relative to a uniform box PDF are increasing with the complexity of the assumed PDF and the anomalies follow a coherent pattern, i.e. the sign of the anomalies in a particular region is the same. The unimodal distributions are more similar among each other than to the double Gaussian PDFs. The difference increases the more CLUBB and EDMF-DualM make use of their two Gaussian PDFs, i.e. the more bimodal they are. One can conclude that the difference between unimodal and bimodal PDFs is more important, than the shape within each mode. The characteristics of the simple unimodal PDFs diagnosed from CLUBB and EDMF-DualM are quite similar, which is not necessarily expected because different input combinations of μ_s and σ_s and different vertical staggering permutations are possible.

Using joint histograms of cloud fraction versus cloud properties, two major specifics of CLUBB and EDMF-DualM were revealed. First, high $\sigma_{q_{l,c}}$ of shallow cumulus clouds shows a distinct second branch in the phase-space of CF and dispersion (Fig. 3.4), while the majority follows the single Gaussian relationship. Second, EDMF-DualM follows primarily the single Gaussian relationship, but as a result of very small variance for low CF, the scaling is more similar to a uniform PDF.

Chapter 4: How important is the subgrid-scale variability of temperature for saturation deficit variability?

The question about the importance of subgrid-scale temperature variability for saturation deficit variability is motivated by the circumstance that most statistical cloud schemes are just formulated in terms of a humidity variability. Therefore the question arises if clouds are really so strongly humidity driven, or what the consequences of neglected temperature variability in terms of cloud properties are. Their joint thermodynamic variation can be studied in the framework of saturation deficit (s). Exploring model simulations using CLUBB and a modified version without subgrid-scale temperature variability, it is shown that subgrid-scale temperature variability can contribute to saturation deficit variability in the same order of magnitude as subgrid-scale humidity variations. The contribution of subgrid-scale temperature variations are a) directly or b) indirectly via its covariability with humidity variations. Within the marine boundary layer the contribution of the covariability term dominates the direct temperature variability influence. Neglecting subgrid-scale temperature variability will lead an underestimate of cloud fraction in the order of few percent when CF is smaller than 0.5 and vice versa to an overestimation for larger CF. However liquid water consistently underestimated. Therefore CF differences that occur between shallow cumulus and stratocumulus region are further amplified. The CF bias is largest in levels of warm marine boundary layer clouds. While the effects attenuates with height for CF, the q_l bias increases further.

Chapter 5: How can global high-resolution satellite observations be used to evaluate subgrid-scale variability produced by GCM cloud parameterizations?

In order to evaluate the subgrid-scale variability of cloud properties produced by GCM cloud parameterizations a novel approach has been explored which is characterized by a) short GCM forecast simulations which provide realistic synoptic situations and b) by comparing the models subgrid variability directly to equivalent observations, that resolve the model's subgrid-scale. MODIS level-2 satellite observations with 1 km nadir horizontal resolution are aggregated onto boxes equivalent to the GCM grid-box size.

In the first part of this section, a spatial comparison of the CLUBB cloud parameterization in the GFDL model and EDMF-DualM cloud parameterization in the ICON model are compared to the MODIS sub-sampled LWP statistics. Some striking differences in their bulk behavior are revealed: Both parameterizations produce extensive stratocumulus cloud decks. However, their spatial extent and LWP-PDF characteristics differ among each other and compared to MODIS. While EDMF-DualM overestimates cloud amount, μ_{LWP} and σ_{LWP}, CLUBB underestimates them compared to MODIS. The in-cloud differences are even more pronounced. Shallow cumulus convection is underestimated by both models, more severely by the EDMF-DualM over land. MODIS frequently observes more organized and deeper cumulus convection embedded in the trade-wind regime, that is just occasionally captured by the EDMF-DualM, while CLUBB produces deeper convective clusters with larger spatial scales than MODIS. The cloud characteristics of a cold air outbreak in the North Atlantic is well captured by both models. Compared to MODIS the associated distribution of the parameterized μ_{LWP} and σ_{LWP} agree well. Eventually the strong synoptic forcing leads to a more realistic simulated boundary layer compared to the less strong forced subtropical marine boundary layer cloud regimes, which is an interesting point for further investigation.

In the second part of this section, a more conceptual comparison of the GCMs and MODIS is done using the instantaneous GCM output. In general the conceptual relationships between LWP and cloud cover are quite well captured by the CLUBB and EDMF-DualM parameterization, however both schemes feature too extreme cloud cover, in particular low CF_{max} and close to 1 CF_{max} occur too frequently. Using these joint histograms the conceptual mechanisms of the cloud parameterization can be highlighted, in particular the cloud regime case switch underlying the EDMF-DualM parameterization (hard-coded at $\mathrm{CF}_{max} = 0.1$) produces unrealistic μ_{LWP} and σ_{LWP} statistics.

Cloud cover is a powerful parameter as it is highly connected to the shape of the LWP-PDF underlying distribution function. For the fully saturated case with $\mathrm{CF}_{max} = 1$, μ_{LWP_c} can theoretically vary independent of σ_{LWP_c}. When $\mathrm{CF}_{max} < 1$, the covariations are restricted by the truncation point of the underlying PDF. Even though CLUBB and EDMF-DualM both use a double Gaussian PDF which could theoretically result in broader spectrum of μ_{LWP_c}-σ_{LWP_c} combinations, they prefer shapes similar to a single Gaussian. In contrast the LWP-PDF statistics from MODIS show a broader range, that cannot be related to simple PDF-types in straightforward manner.

The scale dependence of the LWP-statistics on the sub-sampling domain sizes explored here (100x100 km^2 to 25x25 km^2) is weak. With decreasing domain size, equivalent to a higher GCM spatial resolution, LWP variance, dispersion and skewness are slightly reducing while their relationship to CF_{max} remain.

Concluding, the observed LWP-PDF are more complex than what could be easily explained using simple unimodal and symmetric distribution functions. Still, considering the overall behavior of the CLUBB and EDMF-DualM cloud parameterization in comparison to the cloud schemes used during to fifth climate model intercomparision project (CMIP5), one can conclude that CLUBB and EDMF-DualM show a more physical realistic marine boundary layer structure.

Chapter 6: What can be learned from synoptically realistic cloud resolving models in the context of GCM cloud parameterizations?

First results from the high-resolution ICON-LEM model in context to the HD(CP)^2 project are presented. Within the framework of HD(CP)^2 the ICON model has been extended to perform synoptically forced cloud resolving simulations which produce a realistic atmospheric state. While the ICON-LEM evaluation (Heinze et al., 2016) targets the general performance of boundary layer characteristics and cloud properties, this study focuses on the spatial variability of vertically integrated cloud condensate on GCM-scales explored via sub-sampling the ICON-LEM, the reference model COSMO and MODIS satellite observations. It is found that ICON-LEM is capable to match the satellite observations. Furthermore the different cloud regimes exhibit a distinct signature in terms of their spatial variability. In this context an important aspect are the different input resolutions for. In particular one might assume that variability of ICON-LEM is higher than COSMO because of its finer resolution. However the histogram σ_{CWP} reveals that a) regridding resolutions between 312 m to 2500 m of the original 156 m ICON-LEM does not significantly alter σ_{CWP} of the sub-sampling box and b) the variability produced by ICON-LEM is consistently larger than COSMO and that σ_{CWP} of MODIS is between both models with a tendency towards ICON-LEM.

Additionally the LWP-PDF statistics produced by the GFDL-AM3 CLUBB- and the ICON-GCM EDMF-DualM cloud parameterization are compared in the framework of HD(CP)^2 for the 25 April. Even though this day exhibits the least complex (i.e. most textbook like) boundary layer moist convection case EDMF-DualM has deficiencies to capture the detailed synoptic situation. In contrast, CLUBB is closer to the high-resolution data. The conceptual relationship between the in-cloud LWP-PDF moments and cloud cover reveal a similar scaling of μ_{LWP}, σ_{LWP} and ν_{LWP} with cloud cover of the GCM parameterization compared to the quasi-hemispheric domain of the previous chapter. Particularly the ICON-LEM scaling is quite different, i.e. the in-cloud LWP-PDF moments are larger for low cloud cover. This behavior can be attributed to the smaller but more condensate loaded shallow cumulus clouds compared to MODIS and COSMO. This interesting characteristic is subject of current research and scientific discussion within the HD(CP)^2-community.

7.1. How much complexity is appropriate for a GCM cloud parameterization?

Concluding this question should be split in two aspects: a) the complexity of the parameterized processes that are used to predict the subgrid-scale variability and b) the complexity of the assumed distribution function which determines how the subgrid-scale variability is distributed. Complexity of the parameterized process targets the scope of realism which is incorporated in the formulation of predictive equations, while the complexity of the assumed PDF shape is related to their degrees of freedoms. Certainly these two aspects are deeply connected, since a more elaborate PDF shape requires a more complex set of equations. Generally more degrees of freedoms improve the approximation of an observed cloud scene (Larson et al., 2001a; 2002; Bogenschutz et al., 2010). In particular if multiple cloud regimes occur simultaneously within a grid-box, additional PDF-modes are beneficial since each mode can adopt to a specific task (Perraud et al., 2011). However, when more PDF parameters need to be closed than predictive equations can be inferred from gird-box mean variables, the additional parameters need be closed assuming simple diagnostic relationships, which largely diminishes their capability to respond in a physical manner. Even though diagnostic closures can be redefined and improved (Naumann et al., 2013), this tuning is usually scale-depended, i.e. dependent on the GCM's horizontal resolution.

This study shows that CLUBB and EDMF-DualM subgrid variability is in particular different for low cloud fraction or cloudiness compared to simpler diagnostic assumed PDFs, while for larger cloud fractions the relationship between the saturation deficit PDF and the cloud properties can be well approximated using unimodal PDFs (Fig. 3.15). Compared to the spatial variability of integrated condensate from MODIS observations, the parameterization differ consistently as they are more similar among each other than they are compared to MODIS. In particular the results show that the higher complexity cloud parameterizations do not take the full advantage of their second mode.

To improve the understanding of differences between cloud simulations and observations, it is likely more appropriate to facilitate the use of less complex, unimodal PDFs. Compared to a double Gaussian PDF with 5 degrees of freedom the unimodal PDFs have 2. A few extra degrees of freedom do not seem that much, but what matters for the understanding of the scheme's behavior is the interplay among the saturation deficit PDF moments, i.e the number of covariances ($(N^2 - N)/2$). While unimodal PDFs are characterized by their mean and variability ($(2^2 - 2)/2 = 1$ covariances) a double Gaussian has two means and two variabilities and their respective weight ($(5^2 - 5)/2 = 10$ covariances). Since an intuitive understanding of the multitude of covariances is not straightforward, the use of such complex PDF shapes obfuscates its behavior, which complicates an error attribution and the potential for model improvements. More complex PDFs feature a higher flexibility in terms of possible liquid water skewness, however in this work it is shown that the saturated part of triangular and Gaussian PDFs exhibit significant liquid water skewness too (in contrast to the uniform box PDF). Therefore the use of the triangular or Gaussian PDF is promoted. In particular when the GCM's horizontal resolution keeps increasing the initial weakness of these unimodal PDFs will be diminished.

In case of strong synoptic forcing with heterogeneous large-scale atmospheric conditions, such as a frontal cloud system surrounded by boundary layer clouds, a resolution increase will spatially separate theses regions (Fig. 7.1, left). In synoptically less perturbed cases, radiative-convective equilibrium will alter the large-scale atmospheric conditions which frequently goes along with convective aggregation in regions of strong surface heating (warm sea surface temperatures or cold air advection). The underlying mechanisms are currently in the scientific focus of the cloud-climate community (Tobin et al., 2012; Bretherton and

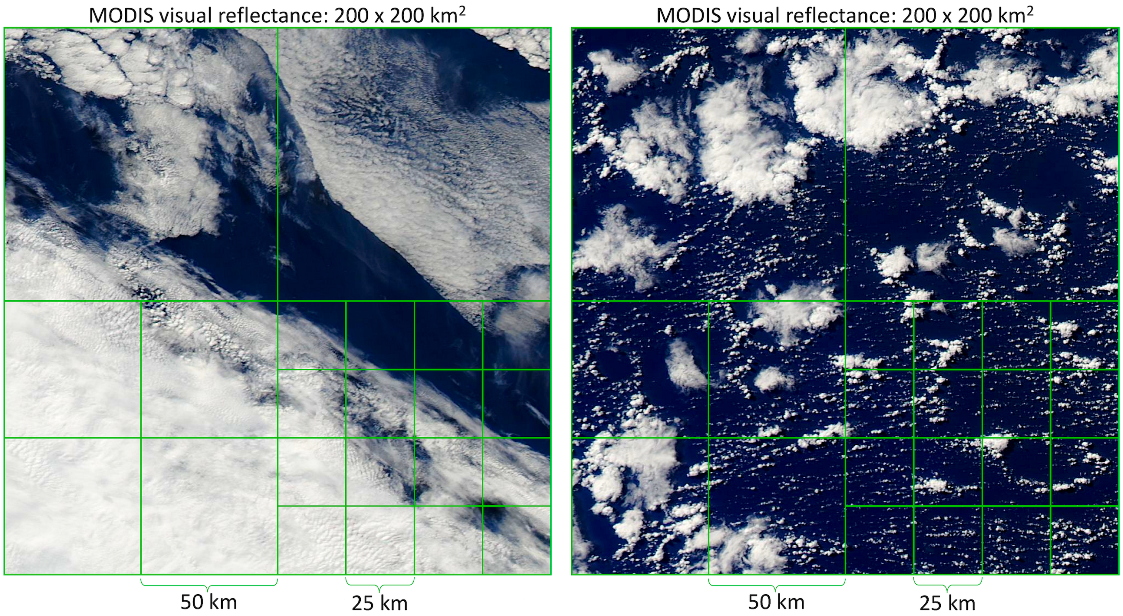


Figure 7.1.: Visible reflectance of Aqua-MODIS at 250 m resolution for 25 April 2013 from worldview.earthdata.nasa.gov showing multiple cloud regimes in heterogeneous large-scale atmospheric conditions. Left: thick frontal cloud in the southern half and shallow boundary layer clouds in the northern half (approx.: 104°E to 102°E, 38°S to 36°S). Right: shallow cumulus and deeper organized convection along the trade-wind trajectory (approx.: 115°E to 113°E, 20°S to 18°S).

Khairoutdinov, 2015; Muller and Bony, 2015; Coppin and Bony, 2015) and are one of the objectives in HD(CP)^2-phase 2. Still the associated horizontal scales are favorable for simpler cloud parameterizations when the spatial resolution of GCMs increase (Fig. 7.1, right). While the convective grey-zone problematic will gain more importance, the parameterization of subgrid-scale cloud properties will benefit. Differentiating between deeper-organized and remaining boundary layer convection on the grid-scale level will reduce the overall liquid water skewness since the liquid water excess relative to the PDF mean will be reduced (since it is partly resolved by the grid-scale), which in turn is in favor of the single Gaussian PDF's performance. In both heterogeneous cases a spatial separation is favorable since the underlying processes are different. A doubling of horizontal resolution can alternatively be interpreted as a quadrupling of cloud parameterization PDF-modes, which are all driven by different grid-box mean states.

When the large-scale atmospheric conditions are homogeneous a resolution increase will not change the performance of unimodal schemes significantly, unless horizontal resolutions reach the kilometer scale when individual shallow cumuli or stratocumulus downdrafts get partially resolved (Fig. 7.2). In the homogeneous case the cloud field is rather self-similar which means that the cloud field exhibits similar spatial characteristics at different scales. Describing such conditions with one PDF mode should be sufficient.

An additional advantage of higher spatial resolutions are improved vertically integrated cloud properties because the cloud overlap problem will largely be reduced once cloud

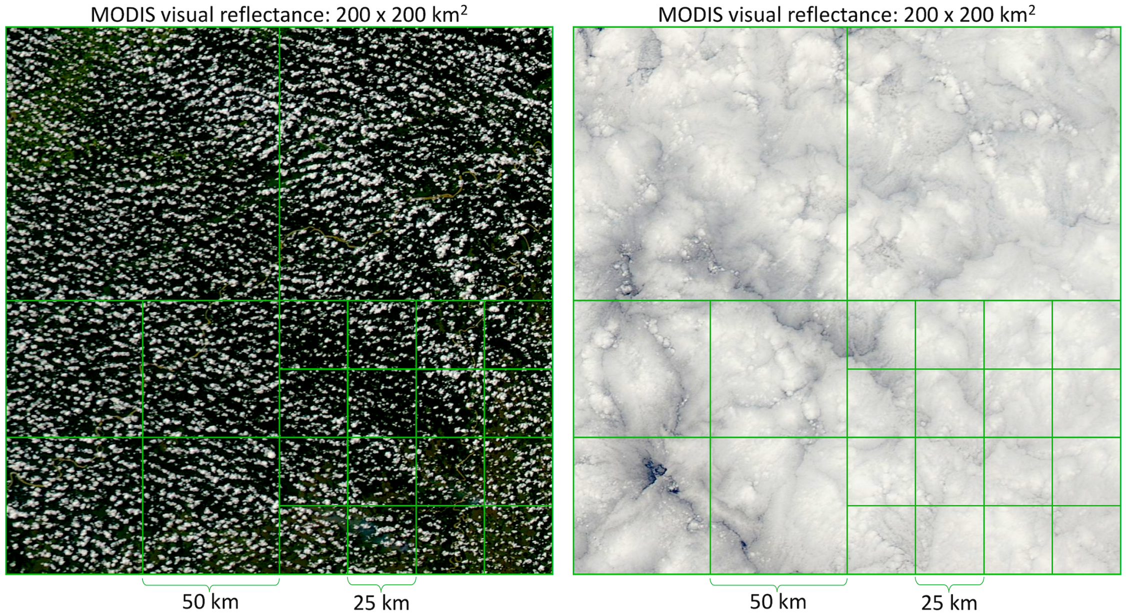


Figure 7.2.: Visible reflectance of Aqua-MODIS at 250 m resolution for 25 April 2013 from worldview.earthdata.nasa.gov showing multiple cloud regimes in homogeneous large-scale atmospheric conditions. Left: shallow cumuli over the western Amazonas (approx.: 69°E to 67°E, 13°S to 11°S) Right: stratocumulus cloud deck (approx.: 84°E to 82°E, 17°S to 15°S).

regimes are separated. For example, a change of the GCM resolution from 200 to 25 km is equivalent to a sub-column generator using 64 columns. Obviously this is still a poor sample for a sub-column generators which usually use up to 10000 sub-columns (Räisänen et al., 2004), but the essential point in this regard is, that the gross distribution is an emerged property of the GCM and not the result of a parameterization, which highlights another aspect: the advanced model capabilities to develop more emergent behavior will not only resolve part of the subgrid-scale variability, but also allow more cloud - circulation interactions which can hardly be parameterized. In this line of reasoning one can conclude that 'solving the correct equations poorly on the grid-scale (i.e. Navier-Stokes equation) might be better than solving the wrong equations well (i.e. subgrid-scale parametrization)' and that 'best cumulus parametrization might be a global model with 5 km spatial resolution' (personal communication with B. Stevens, 2016).

Summarizing, the suggestion of this work is to use less complex assumed PDFs for cloud parameterizations in operational GCMs. Accordingly one has to recognize that this decision goes along with a trade-off in realism in cases of multi-modal cloud regimes. However the reduced complexity allows to attribute the shortcomings of the parameterization's behavior in a straightforward and more intuitive manner. If the range of uncertainty induced by the subgrid-scales is better established, the grid-scale shortcomings can be explored more consistently since the ambiguity of grid-scale and subgrid-scale errors can be reduced. The mentioned shortcomings of less complex PDFs can be diminished when the model's spatial resolution increases, which is already the trend in the model development. Lastly there is a practical, non scientific reason that supports less complex cloud parameterizations (or parameterizations in general): from personal experience the original implementation is drifting over time to fulfill the needs of the hosting model. This deviation induces practical problems when the fluctuation of scientists, responsible for the model development, is high, since the knowledge about the implications of ad-hoc tweaks is largely attributed to them.

Nevertheless the advancement of complex and more physical cloud parameterizations is essential since they can be considered as the frontier of the community's understanding on a specific subgrid-scale problem. In particular they are needed to explain cloud observations or cloud resolving simulations from a conceptual point of view, which in turn improves the understanding of involved physical processes.

7.2. Outlook

The first results from the HD(CP)^2 cloud resolving modeling effort are promising and will be followed up by extending the analysis of Chapter 6 to the upcoming next simulation days, that cover an even wider range of synoptic situations. In particular the 3D condensate distributions will be explored and the analytical relationships between 2D and 3D fields (established in Chapter 2) generalized. For this purpose new post-processing tools need to be developed in order to process the data amount of $\mathcal{O}(50\text{ TB})$ per simulation day.

Within the second phase of the HD(CP)^2 project (start April 2016) high-resolution simulations over the tropical Atlantic on even larger domains will be performed. Larger domains and more unperturbed conditions will facilitate the inter-comparison of the high-resolution simulations, MODIS observations and the GCM cloud parameterization results, because synoptic-scale features can be fully captured which allows a better qualitative comparison. The spatial patterns of subgrid-scale integrated cloud condensate variability can be geographically referenced on the basis of cloud regimes analog to Chapter 5. This comparison can furthermore be quantified by applying an objective cloud regime identification on the basis of instantaneous cloud properties, e.g. using neural network clustering techniques such as Wood and Hartmann (2006); Muhlbauer et al. (2014).

A further promising idea to reveal the spatial scale dependence of unimodal and bimodal PDF's performance. This can be achieved by coupling a range of assumed PDFs directly to the high-resolution model output (Rosch et al., 2015) by using the spatially aggregated values as input. This approach would provide a consistent comparison of parametrized subgrid-scale- and resolved cloud variability. In that manner different sub-sampling box sizes (aggregation scales) can be tested to quantify the scale dependence and explore the relationships between the PDF types and their associated cloud properties. The coupling of unimodal PDF to the aggregated high-resolution output can be done as in Chapter 3, while the coupling of bimodal PDF is slightly more complex (Larson et al., 2001a; Naumann et al., 2013).

Besides the outlined positive aspects of high-resolution modeling framework, a general downside is that the model evolution shifted towards technical developments and highly optimized source codes which make the implementation of own ideas more challenging. Additionally the deeply routed high performance libraries, the increasing hardware demands and the multitude of auxiliary model setup data require evermore effort, which underlines the importance for an improved workflow. One promising approach to tackle the large output amounts is to advance online diagnostics, i.e. deriving post-processed information during model runtime. Of particular interest to the topic of this work is the joint PDF diagnostic of cloud parameterization related parameters. In this regard, the influence of the aggregation scale, the number of PDF-bins or the joint variables combinations can be further explored.

Given a finite amount of computing resources, in terms of allocated computing time on an shared infrastructure or in terms of maximum available stack memory, a trade-off mainly between horizontal resolution and domain size has to be made. For high-resolution simulations which should cover vast domains, it might be beneficial to compromise on the horizontal resolution, gaining a speed-up due to fewer grid-boxes and longer timesteps. Usually large-eddy simulations (like the HD(CP)^2 simulations in Chapter 6) do not feature fractional cloudiness, however spatial resolutions in the order of one kilometer require the incorporation of subgrid-scale cloud variability. In this sense, there is an urgency to develop, explore and evaluate a new class of high-resolution cloud parameterizations.

APPENDIX

A. Theoretical relationships

In addition to the in-cloud liquid water PDF (q_l-PDF) scaling in relation to cloud fraction (CF) (shown in Chapter 2), these relationships are now shown for the grid-box scale for the idealized single (Fig. A.3) and double (Fig. A.4) layer cloud cloud case.

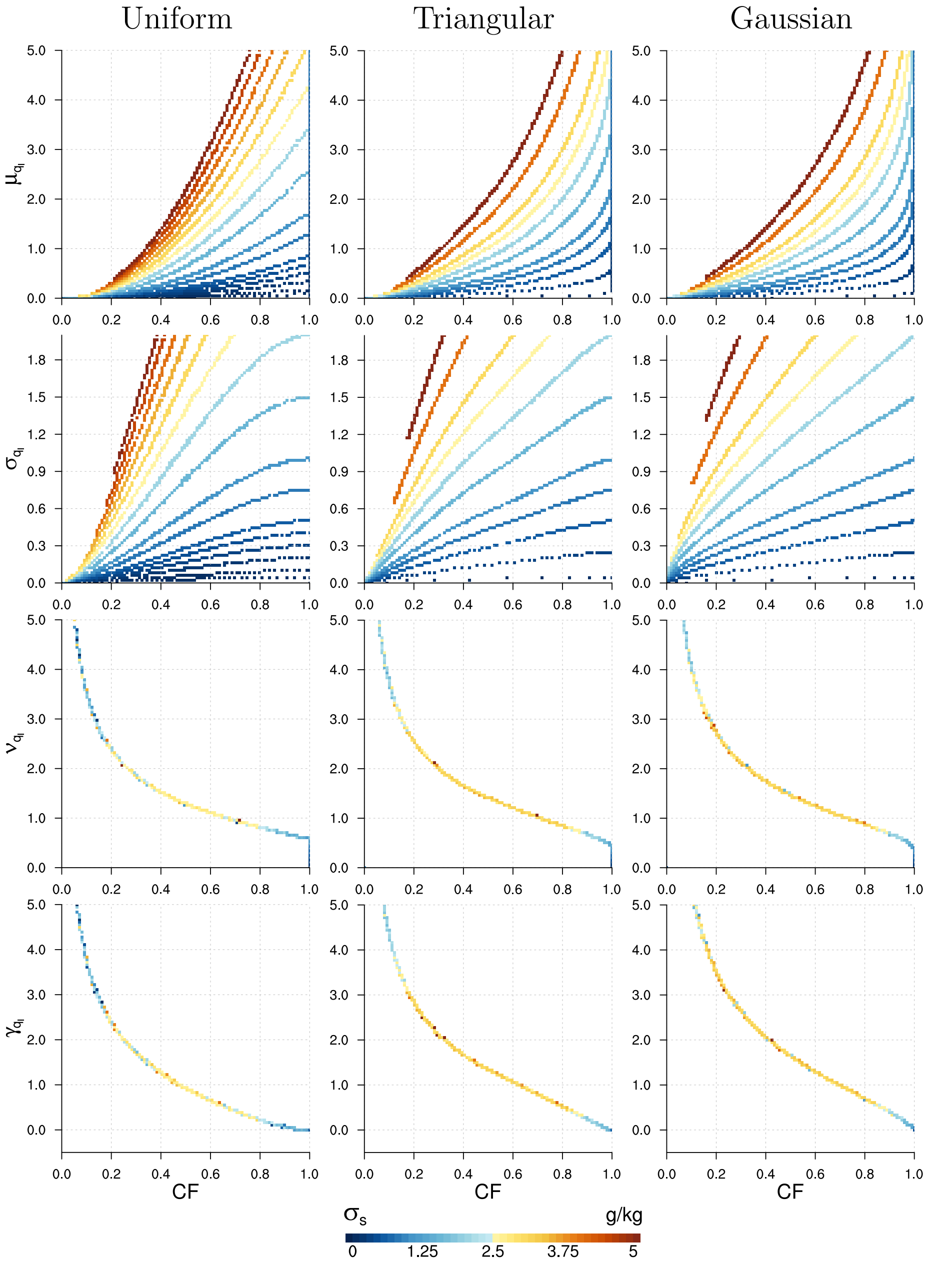


Figure A.3.: Scaling of liquid water q_l-PDF parameters in relation to cloud fraction (CF) within one layer. The colors indicate the saturation deficit variability (σ_s). The q_l-PDF is computed across a μ_s-range from -5 to 5 with an discretization of 500 steps and a multitude of σ_s. For each μ_s-σ_s combination a Monte-Carlo methods (N=100000 values) is used to diagnose the cloud properties. The μ_s-$\sigma_s$phase-space is subsequently rearranged in relation to CF. Complementary to Fig. 2.4

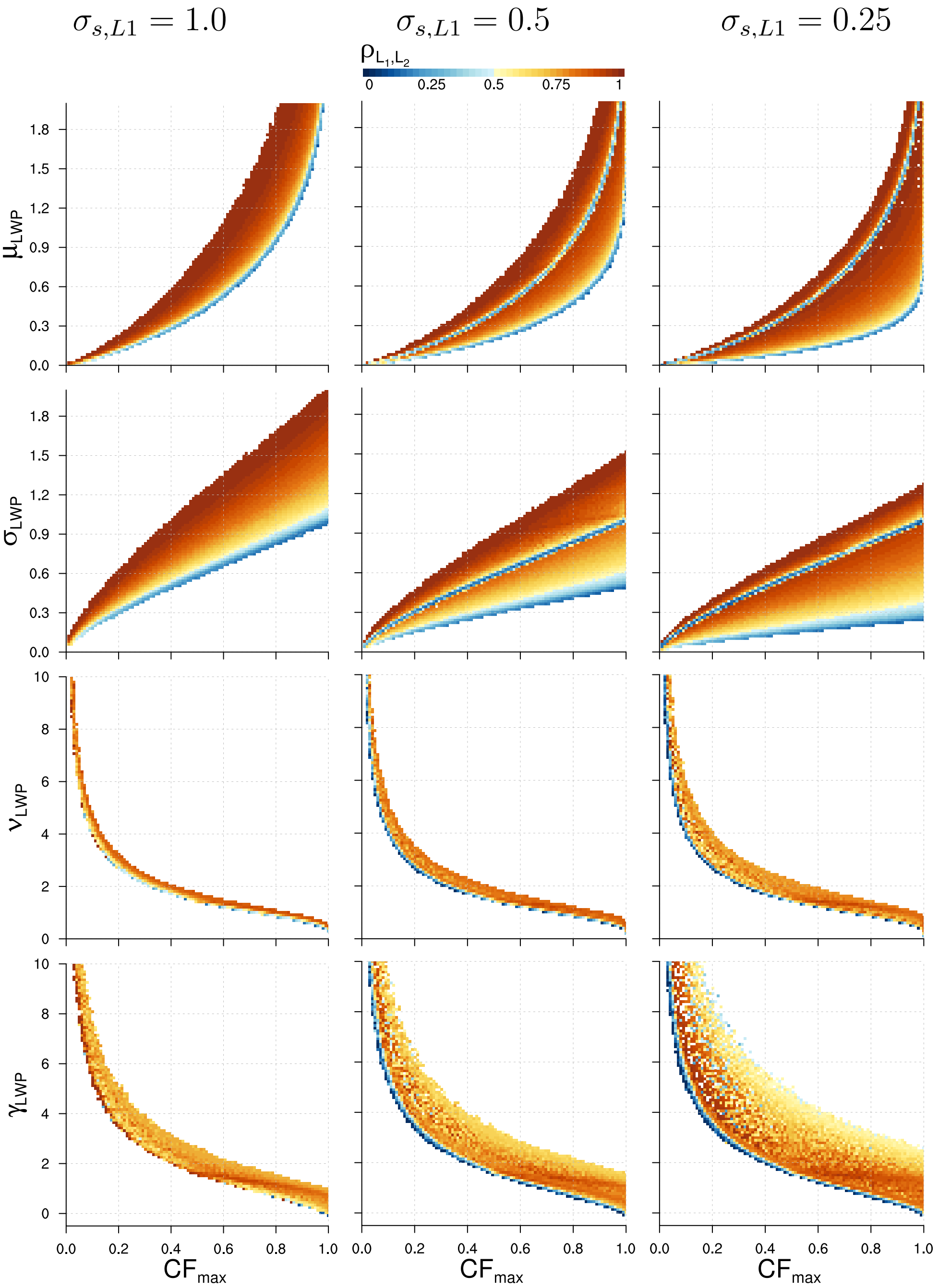


Figure A.4.: Monte-Carlo simulations: scaling of liquid water LWP-PDF parameters relative to cloud cover (CF_{max}) derived from a two layer cloud. The colors indicate the inter-layer correlation ($\rho_{1,2}$). The LWP-PDF is diagnosed from three s-PDF combinations with the same $\sigma_{s,2} = 1$ and different $\sigma_{s,1} = 1, 0.5, 0.25$ (left to right column). Data rearranged to show the relationships of CF_{max} versus μ_{LWP_c} (row 1), σ_{LWP_c} (row 2), ν_{LWP_c} (row 3), γ_{LWP_c} (row 4). Complementary to Fig. 2.9

In addition to Figs. 2.9 and A.4 the scaling of grid-box (Fig. A.5) and in-cloud (Fig. A.5) vertically integrated q_l-moments, i.e. liquid water path (LWP), are shown in terms of terms of $\mathrm{CF}_{\mathrm{Layer\ 1}}$ and $\mathrm{CF}_{\mathrm{Layer\ 2}}$.

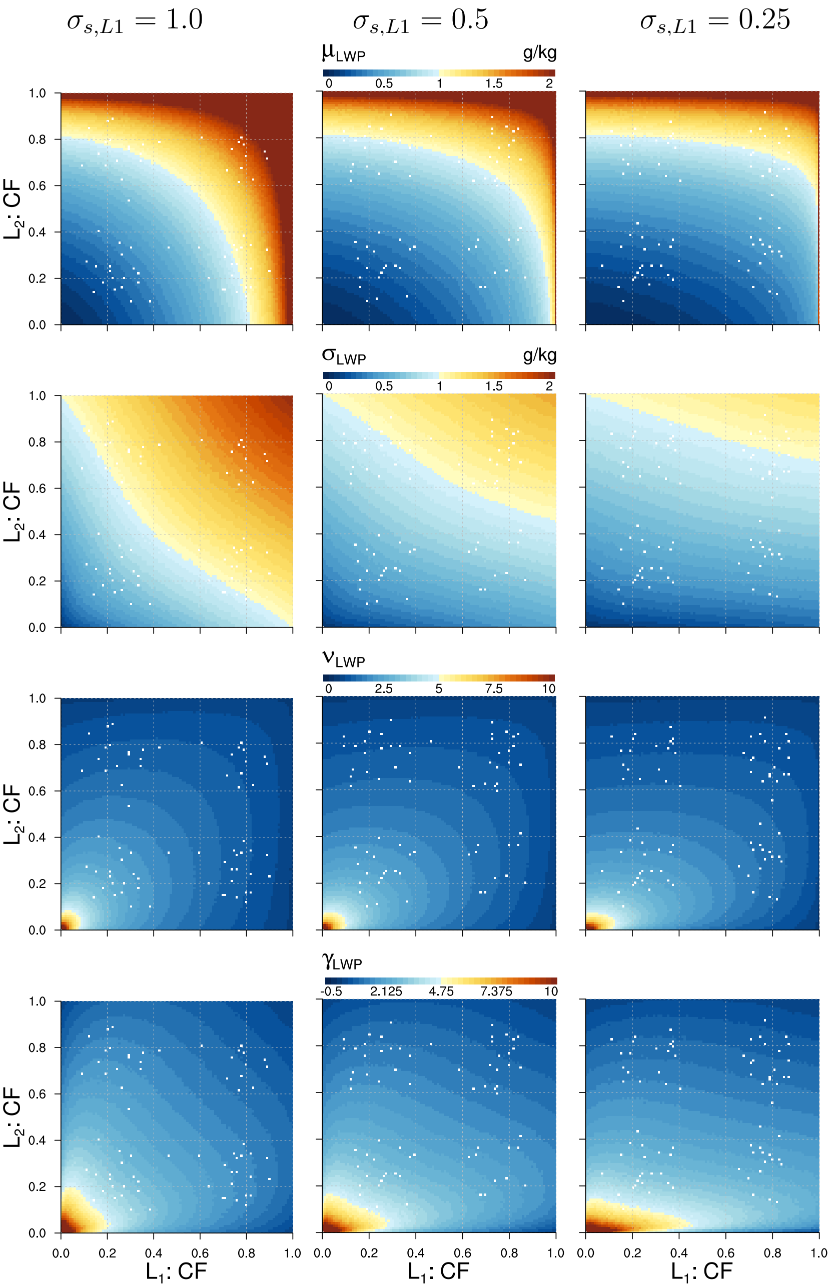


Figure A.5.: Monte-Carlo simulations: scaling of grid-box liquid water LWP-PDF parameters (color) relative to $\mathrm{CF}_{\mathrm{Layer\ 1}}$ and $\mathrm{CF}_{\mathrm{Layer\ 2}}$ derived from a two layer cloud. The LWP-PDF is diagnosed from three s-PDF combinations with the same $\sigma_{s,2} = 1$ and different $\sigma_{s,1} = 1, 0.5, 0.25$ (left to right column). μ_{LWP} (row 1), σ_{LWP} (row 2), ν_{LWP} (row 3), γ_{LWP} (row 4).

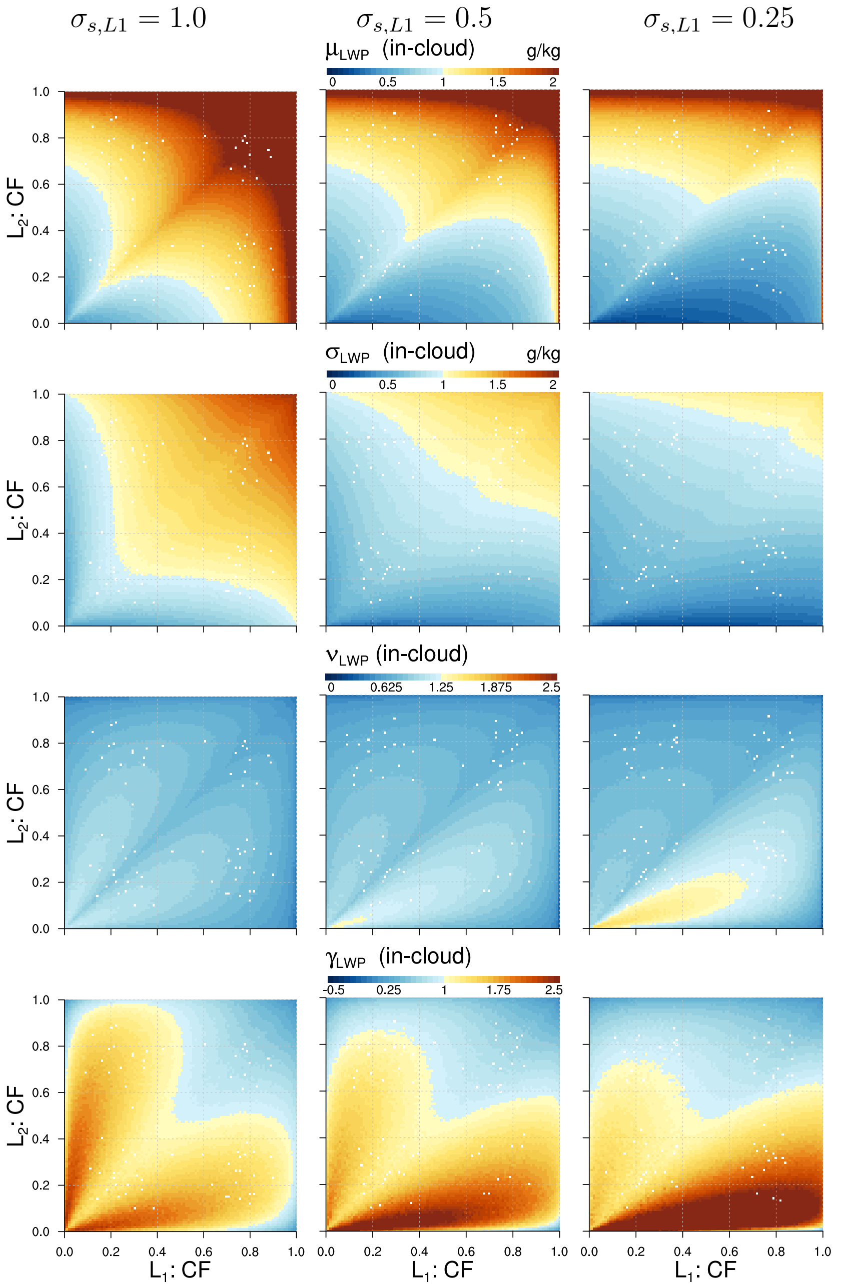


Figure A.6.: Monte-Carlo simulations: scaling of in-cloud liquid water LWP-PDF parameters (color) relative to $\mathrm{CF}_{\mathrm{Layer\ 1}}$ and $\mathrm{CF}_{\mathrm{Layer\ 2}}$ derived from a two layer cloud. The LWP-PDF is diagnosed from three s-PDF combinations with the same $\sigma_{s,2} = 1$ and different $\sigma_{s,1} = 1, 0.5, 0.25$ (left to right column). μ_{LWP_c} (row 1), σ_{LWP_c} (row 2), ν_{LWP_c} (row 3), γ_{LWP_c} (row 4).

B. Subgrid-scale variability from PDF cloud schemes in GCMs

In addition to the anomalies of vertically integrated grid-cloud liquid water path (LWP) characteristics of different complexity assumed PDFs (Chapter 3), they are now shown for for the in-cloud liquid water path (LWP). The characteristics derived from GFDL-AM3 CLUBB are in Fig. B.7 and the ones from the ICON-GCM EDMF-DualM in Fig. B.8). In the same manner, the vertical histograms of the grid-box scale q_l-PDF moments are shown in Fig. B.9 (corresponding the in-cloud version Fig. 3.14) and the joint histograms of cloud fraction (CF) vs. in-cloud q_l-PDF moments are shown in Fig. B.10 (corresponding the in-cloud version Fig. 3.15).

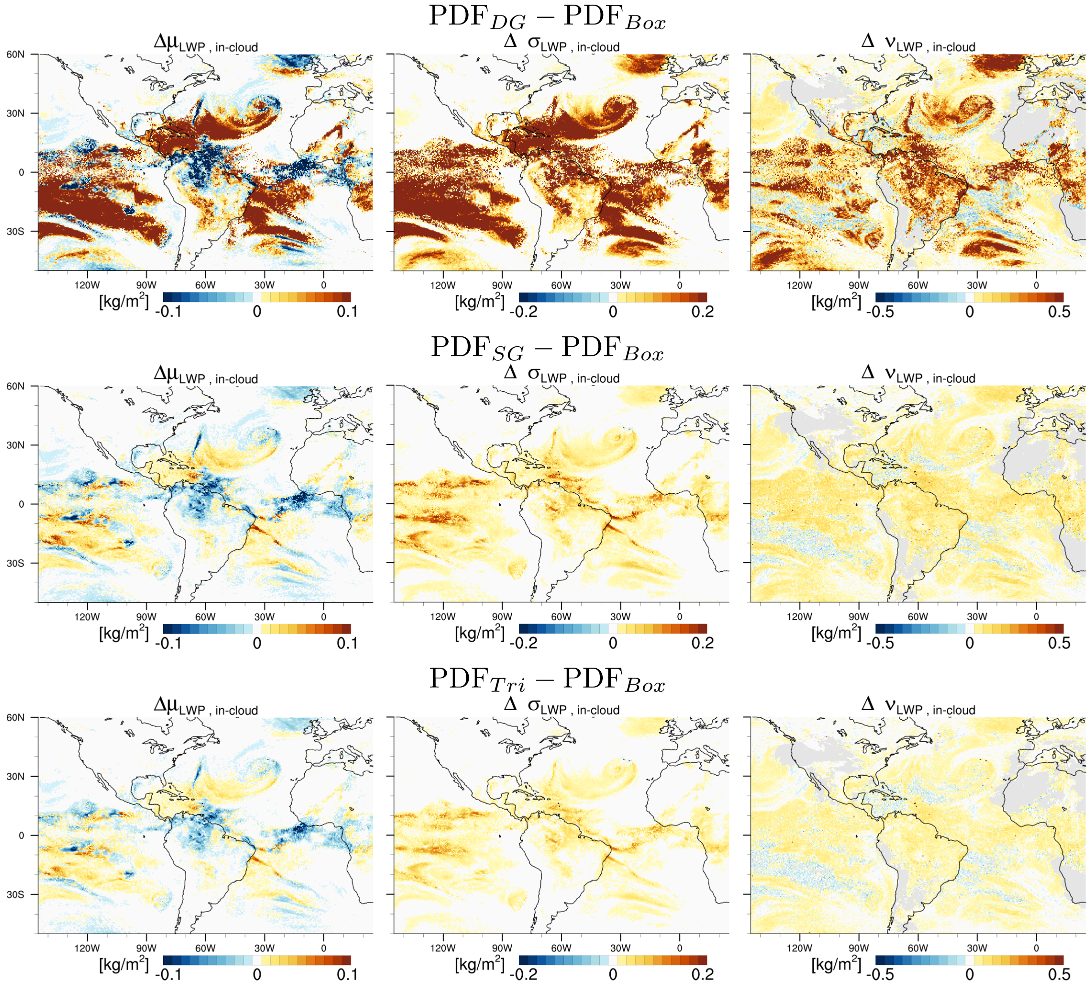


Figure B.7.: GFDL-AM3 CLUBB, setup equivalent to Fig. 3.1: anomalies of vertically integrated in-cloud liquid water path (LWP) characteristics, i.e. difference between the original double Gaussian (DG) and the uniform (box) PDF (top row), difference between the single Gaussian (SG) and the box PDF (middle row) and difference between the triangular (Tri) and box PDF (bottom row). Column 1: in-cloud LWP mean difference $\Delta\mu_{\mathrm{LWP}_c}$. Column 2: in-cloud LWP standard deviation difference $\Delta\sigma_{\mathrm{LWP}_c}$. Column 3: in-cloud LWP dispersion difference $\Delta\nu_{\mathrm{LWP}_c}$. Grey color indicates cloud free regions, i.e. missing values.

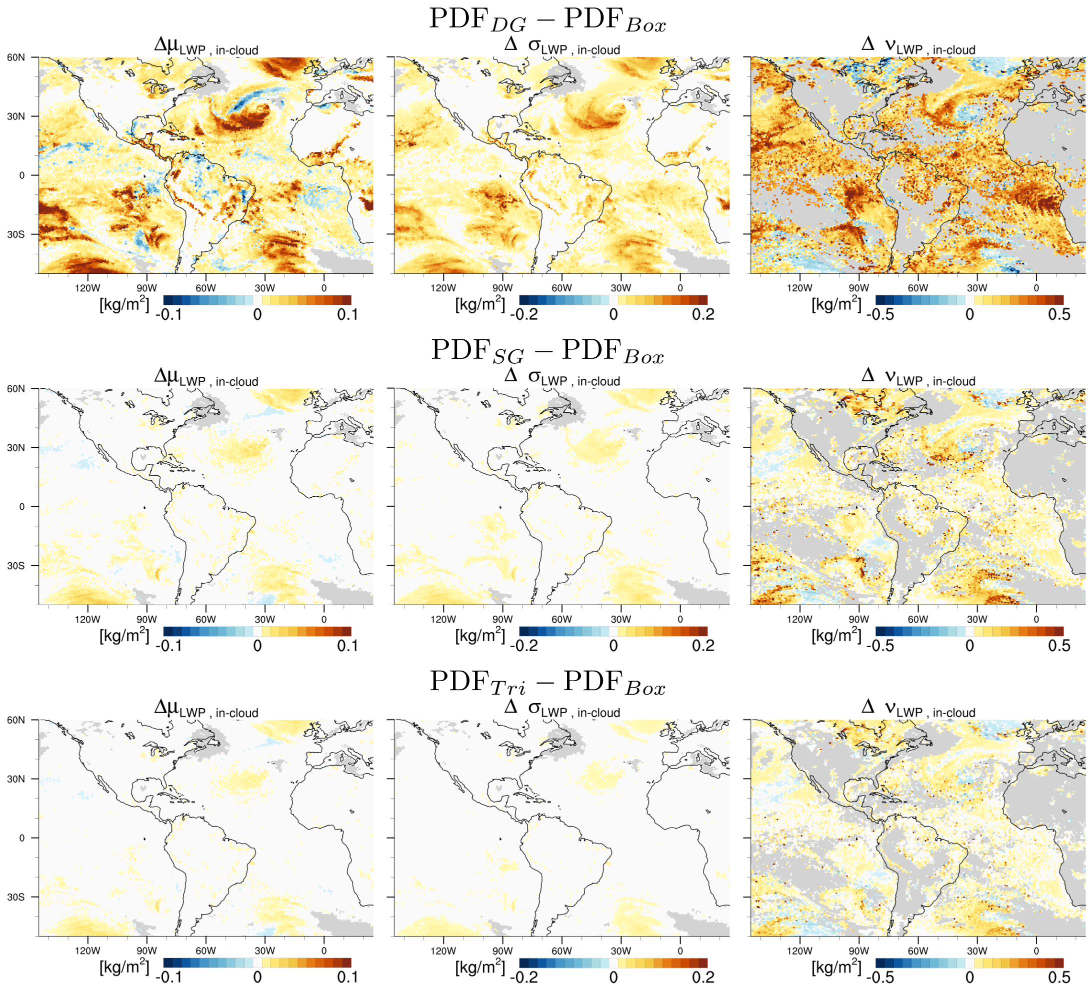


Figure B.8.: ICON EDMF-DualM, setup equivalent to Fig. 3.7: anomalies of vertically integrated in-cloud liquid water path (LWP) characteristics, i.e. difference between the original double Gaussian (DG) and the uniform (box) PDF (top row), difference between the single Gaussian (SG) and the box PDF (middle row) and difference between the triangular (Tri) and box PDF (bottom row). Column 1: in-cloud LWP mean difference $\Delta\mu_{\mathrm{LWP}_c}$. Column 2: in-cloud LWP standard deviation difference $\Delta\sigma_{\mathrm{LWP}_c}$. Column 3: in-cloud LWP dispersion difference $\Delta\nu_{\mathrm{LWP}_c}$. Columns with no EDMF-DualM activity are set to missing value (gray).

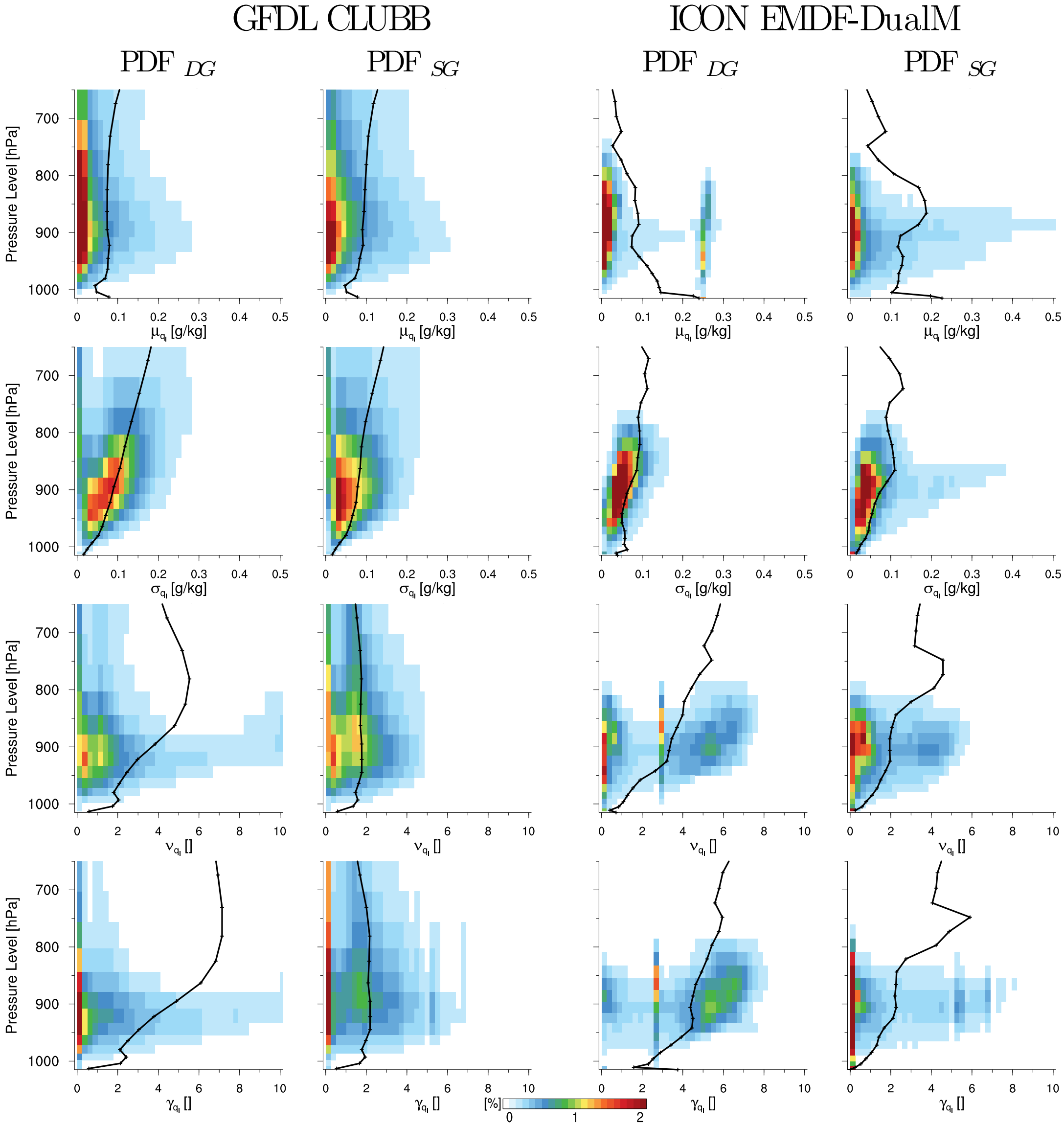


Figure B.9.: Vertical histograms analog to Fig. 3.14, now for the grid-box scale of liquid water q_l-PDF. Row 1: q_l-mean (μ_{q_l}). Row 2: q_l-standard deviation (σ_{q_l}). Row 3: q_l-dispersion (ν_{q_l}). Row 4: q_l-skewness (γ_{q_l}).

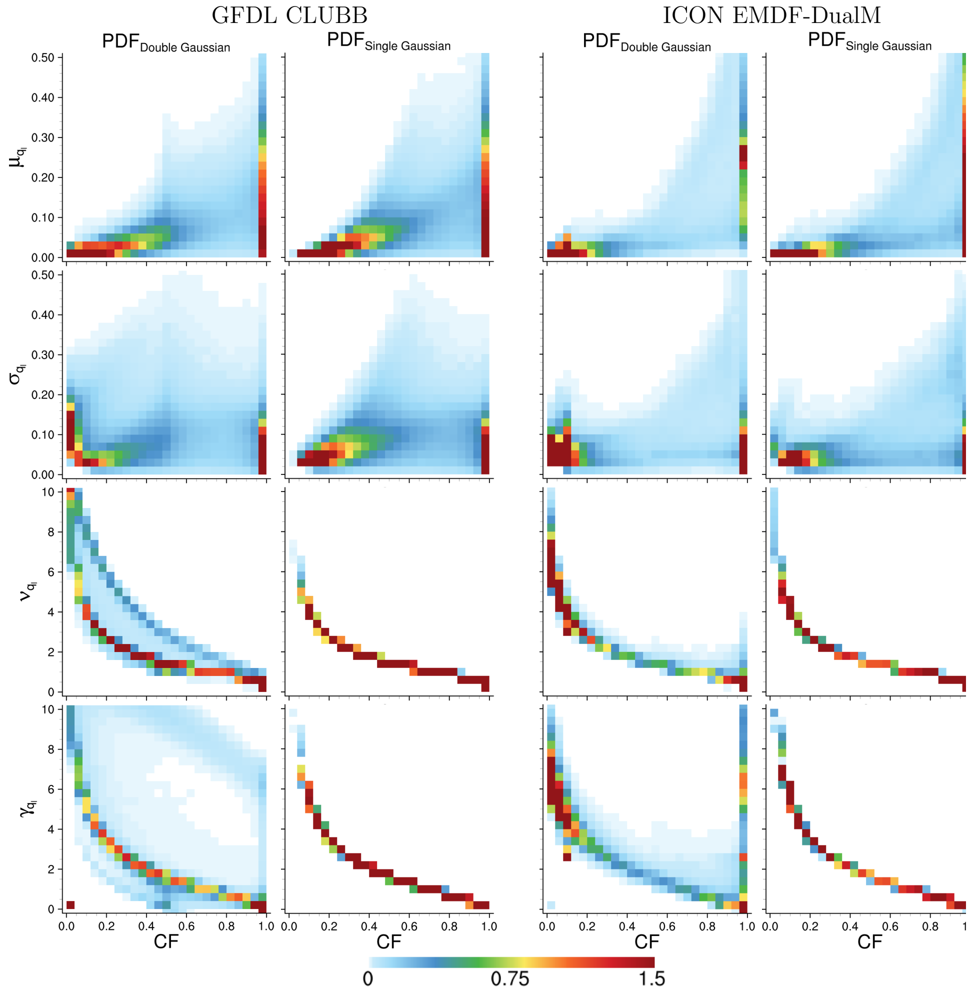


Figure B.10.: Joint histograms of cloud fraction (CF) vs. liquid water q_l-PDF moments analog to Fig. 3.15. Row 1: q_l-mean (μ_{q_l}). Row 2: q_l-standard deviation (σ_{q_l}). Row 3: q_l-dispersion (ν_{q_l}). Row 4: q_l-skewness (γ_{q_l}).

C. Importance of subgrid temperature variability for cloud parameterization

In Chapter 4 the GFDL-AM3 CLUBB simulation are performed for a 30 day period. Fig. C.11 shows spatial distribution of saturation deficit variability (σ_s), cloud fraction (CF), and in-cloud liquid water $\mu_{q_{l,c}}$, which is to some extent redundant to Figs. 3.1, 3.2, but summarizes the distribution of subgrid-scale variability without the influence of synoptic events.

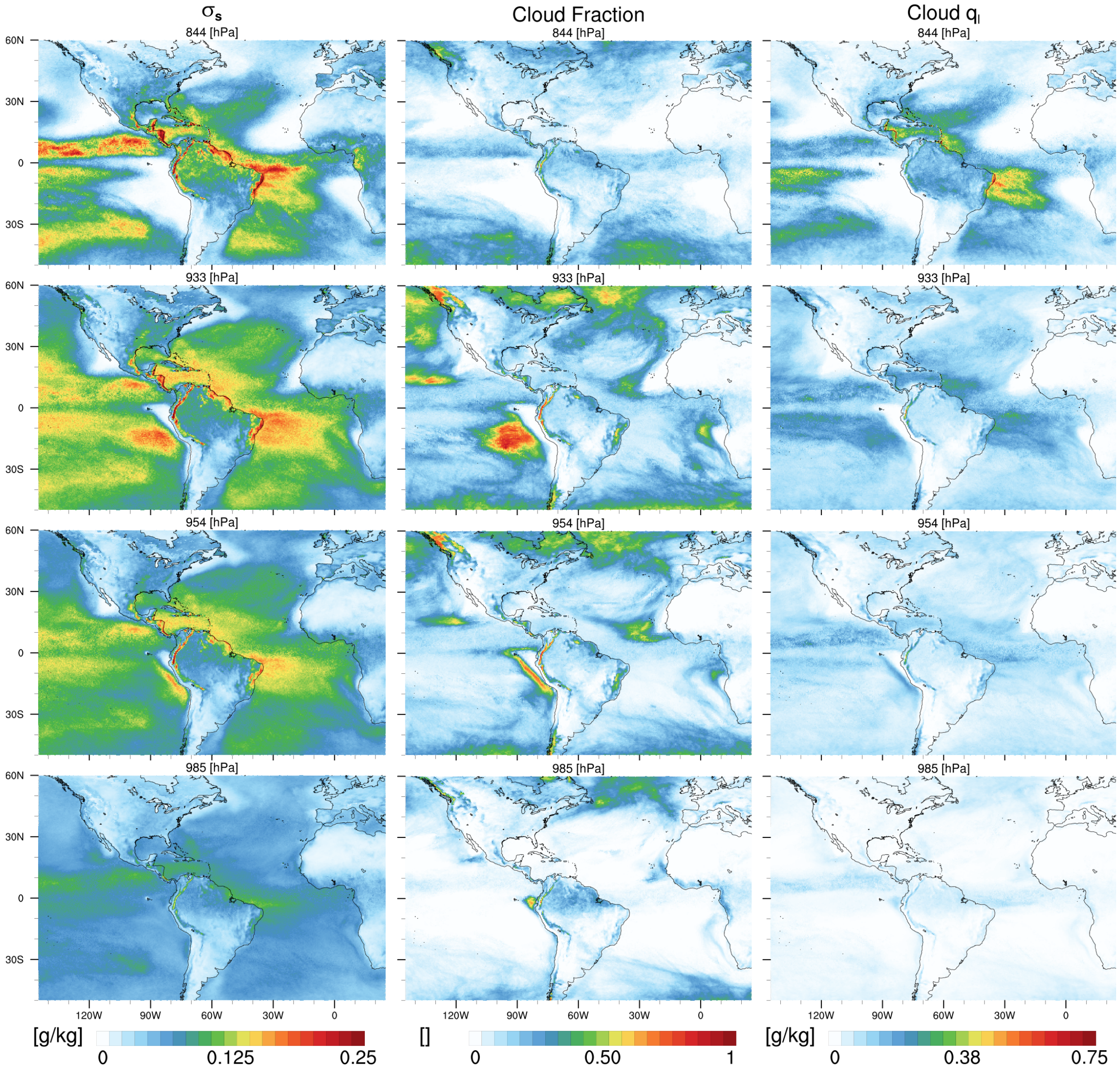


Figure C.11.: Spatial distribution of saturation deficit variability σ_s (left), cloud fraction CF (middle), and in-cloud liquid water $\mu_{q_{l,c}}$ (right) in four height layers most relevant to warm boundary layer clouds. Derived from instantaneous model output of a GFDL-AM3 CLUBB AMIP simulation for 30d in May starting from NCEP nudged base run.

D. Large-scale evaluation

For consistency and the joint histograms of cloud cover (CF_{max}) vs. grid-box LWP-PDF moments are shown in Fig. D.12 (corresponding the in-cloud version Fig. 5.8 in Chapter 5). Note that grid-scale variability is largely driven by cloud fraction or cloud cover, therefore it is more appropriate to study in the in-cloud relationships in context the sub-grid scale evaluation.

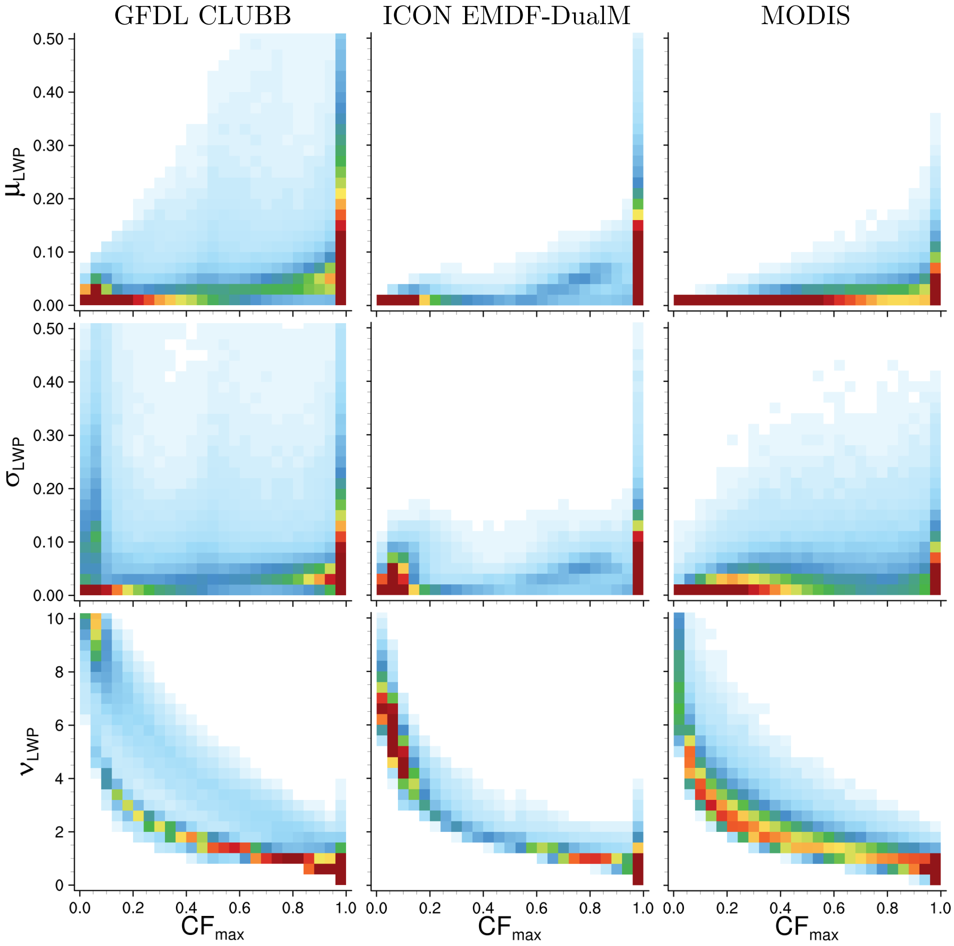


Figure D.12.: Joint histograms showing the relationships between the grid-box LWP-PDF moments and the cloud cover (CF_{max}). Instantaneous GCM output and all grid-boxes and MODIS sub-sampling domains for the large-scale domain (145°W to 25°E, 50°S to 60°N) for the 25. Apr. 2013 are used. Color show the frequency of occurrence (FoO): white 0%, blue ~ 0.5%, green ~ 1.0%, red ~ 1.5%.

E. Condensate variability in context to HD(CP)^2

For consistency and the joint histograms of cloud cover (CF_{max}) vs. grid-box LWP-PDF moments are shown in Fig. E.13 (corresponding the in-cloud version Fig. 6.5 in Chapter 6).

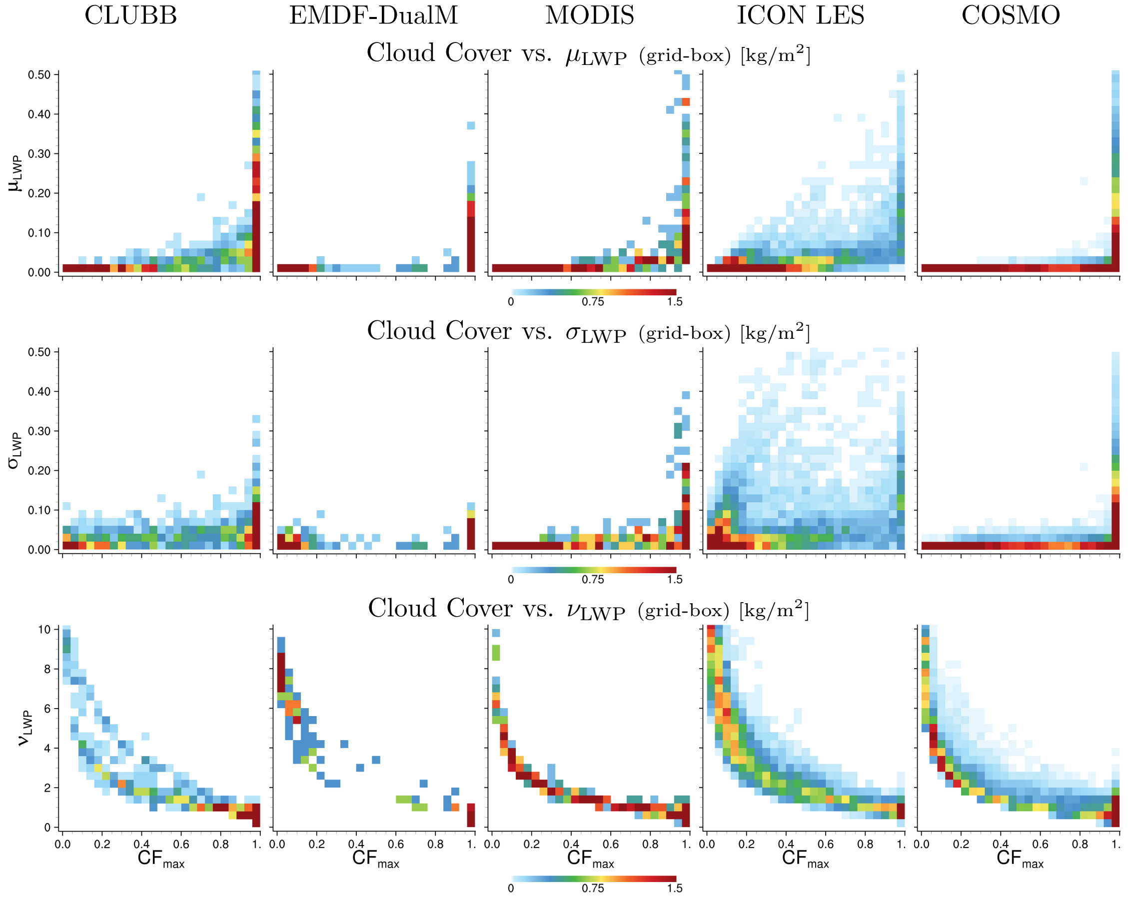


Figure E.13.: Joint histograms of cloud fraction (CF) vs. liquid water path LWP-PDF moments analog to Fig. 6.4 but using instantaneous data. Row 1: LWP-mean (μ_{LWP}). Row 2: LWP-standard deviation (σ_{LWP}). Row 3: LWP-dispersion (ν_{LWP}). Row 4: LWP-skewness (γ_{LWP}).

BIBLIOGRAPHY

Arakawa, A., 1975: *Modelling clouds and cloud processes for use in climate model.* Global Atmospheric Research Program Publication Series, no. 16 (ICSU/WMO). The Physical Basis of Climate and Climate Modeling, 183–197.

Arakawa, A., 2004: The cumulus parameterization problem: Past, present, and future. *Journal of Climate*, **17 (13)**, 2493–2525.

Arakawa, A. and J.-H. Jung, 2011: Multiscale modeling of the moist-convective atmospherea review. *Atmospheric Research*, **102 (3)**, 263–285.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. *Monthly Weather Review*, **139 (12)**, 3887–3905.

Barker, H., R. Pincus, and J. Morcrette, 2002: The monte carlo independent column approximation: Application within large-scale models. *Proceedings of the GCSS/ARM workshop on the representation of cloud systems in large-scale models.*

Barker, H. and P. Räisänen, 2004: Neglect by GCMs of subgrid-scale horizontal variations in cloud-droplet effective radius: A diagnostic radiative analysis. *Quarterly Journal of the Royal Meteorological Society*, **130 (600)**, 1905–1920.

Barker, H., et al., 2003: Assessing 1D atmospheric solar radiative transfer models: Interpretation and handling of unresolved clouds. *Journal of Climate*, **16 (16)**, 2676–2699.

Barker, H. W. and B. A. Wielicki, 1997: Parameterizing grid-averaged longwave fluxes for inhomogeneous marine boundary layer clouds. *Journal of the atmospheric sciences*, **54 (24)**, 2785–2798.

Barker, H. W., B. A. Wiellicki, and L. Parker, 1996: A parameterization for computing grid-averaged solar fluxes for inhomogeneous marine boundary layer clouds. part II: Validation using satellite data. *Journal of the atmospheric sciences*, **53 (16)**, 2304–2316.

Barthlott, C. and C. Hoose, 2015: Spatial and temporal variability of clouds and precipitation over germany: multiscale simulations across the" gray zone". *Atmospheric Chemistry & Physics Discussions*, **15 (12)**.

Bechtold, P., M. Köhler, T. Jung, F. Doblas-Reyes, M. Leutbecher, M. J. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time-scales. *Quarterly Journal of the Royal Meteorological Society*, **134 (634)**, 1337–1351.

Blossey, P. N., et al., 2013: Marine low cloud sensitivity to an idealized climate change: The CGILS LES intercomparison. *Journal of Advances in Modeling Earth Systems*, **5 (2)**, 234–258.

Bogenschutz, P. A., A. Gettelman, H. Morrison, V. E. Larson, C. Craig, and D. P. Schanen, 2013: Higher-order turbulence closure and its impact on climate simulations in the community atmosphere model. *Journal of Climate*, **26 (23)**, 9655–9676.

Bogenschutz, P. A., S. K. Krueger, and M. Khairoutdinov, 2010: Assumed probability density functions for shallow and deep convection. *Journal of Advances in Modeling Earth Systems*, **2 (4)**.

Bony, S. and J. Dufresne, 2005: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophys. Res. Lett.*, **32**, L20 806, doi:DO:10.1029/2005GL023851.

Bony, S., et al., 2015: Clouds, circulation and climate sensitivity. *Nature Geoscience*, **8 (4)**, 261–268.

Boucher, O., et al., 2013a: Chapter 7: Clouds and aerosols. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Boucher, O., et al., 2013b: Clouds and aerosols. *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*, Cambridge University Press, 571–657.

Bougeault, P., 1981a: Modeling the trade-wind cumulus boundary layer. part I: Testing the ensemble cloud relations against numerical data. *Journal of the Atmospheric Sciences*, **38 (11)**, 2414–2428.

Bougeault, P., 1981b: Modeling the trade-wind cumulus boundary layer. part II: A high-order one-dimensional model. *Journal of the Atmospheric Sciences*, **38 (11)**, 2429–2439.

Boutle, I., S. Abel, P. Hill, and C. Morcrette, 2014a: Spatial variability of liquid cloud and rain: Observations and microphysical effects. *Quarterly Journal of the Royal Meteorological Society*, **140 (679)**, 583–594.

Boutle, I., J. Eyre, and A. Lock, 2014b: Seamless stratocumulus simulation across the turbulent gray zone. *Monthly Weather Review*, **142 (4)**, 1655–1668.

Bretherton, C. S. and M. Khairoutdinov, 2015: Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *Journal of Advances in Modeling Earth Systems*.

Brient, F. and S. Bony, 2013: Interpretation of the positive low-cloud feedback predicted by a climate model under global warming. *Climate Dynamics*, **40 (9-10)**, 2415–2431.

Brient, F., T. Schneider, Z. Tan, S. Bony, X. Qu, and A. Hall, 2015: Shallowness of tropical low clouds as a predictor of climate models response to warming. *Climate Dynamics*, 1–17.

Brueck, M., L. Nuijens, and B. Stevens, 2015: On the seasonal and synoptic time scale variability of the north atlantic trades and its low-level clouds. *Journal of the Atmospheric Sciences*, **72 (4)**, 1428 – 1446, doi:http://dx.doi.org/10.1175/JAS-D-14-0054.1.

Cahalan, R. F., W. Ridgway, W. J. Wiscombe, T. L. Bell, and J. B. Snider, 1994: The albedo of fractal stratocumulus clouds. *Journal of the Atmospheric Sciences*, **51 (16)**, 2434–2455.

Chen, J.-M., 1991: Turbulence-scale condensation parameterization. *Journal of the atmospheric sciences*, **48 (12)**, 1510–1512.

Cheng, A. and K.-M. Xu, 2009: A PDF-based microphysics parameterization for simulation of drizzling boundary layer clouds. *Journal of the Atmospheric Sciences*, **66 (8)**, 2317–2334.

Considine, G., J. A. Curry, and B. Wielicki, 1997: Modeling cloud fraction and horizontal variability in marine boundary layer clouds. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **102 (D12)**, 13 517–13 525.

Coppin, D. and S. Bony, 2015: Physical mechanisms controlling the initiation of convective self-aggregation in a general circulation model. *Journal of Advances in Modeling Earth Systems*, **7 (4)**, 2060–2078.

de Roode, S. R. and P. G. Duynkerke, 1996: Dynamics of cumulus rising into stratocumulus as observed during the first lagrangianexperiment of ASTEX. *Quarterly Journal of the Royal Meteorological Society*, **122 (535)**, 1597–1623.

Dipankar, A., B. Stevens, R. Heinze, C. Moseley, G. Zängl, M. Giorgetta, and S. Brdar, 2015: Large eddy simulation using the general circulation model ICON. *Journal of Advances in Modeling Earth Systems*, **7 (3)**, 963–986.

Doms, G., et al., 2011: A description of the nonhydrostatic regional COSMO model, part II: Physical parameterization. Tech. rep., Deutscher Wetterdienst: Offenbach, Germany (accessed 9 May 2016).

Donner, L. J., C. J. Seman, R. S. Hemler, and S. Fan, 2001: A cumulus parameterization including mass fluxes, convective vertical velocities, and mesoscale effects: Thermodynamic and hydrological aspects in a general circulation model. *Journal of climate*, **14 (16)**, 3444–3463.

Donner, L. J., et al., 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *Journal of Climate*, **24 (13)**.

Dorrestijn, J., D. T. Crommelin, A. P. Siebesma, H. Jonker, and C. Jakob, 2015: Stochastic parameterization of convective area fractions with a multicloud model inferred from observational data. *Journal of the Atmospheric Sciences*, **72 (2)**, 854–869.

Dorrestijn, J., D. T. Crommelin, A. P. Siebesma, and H. J. Jonker, 2013: Stochastic parameterization of shallow cumulus convection estimated from high-resolution model data. *Theoretical and Computational Fluid Dynamics*, **27 (1-2)**, 133–148.

Emanuel, K. A., 1994: *Atmospheric convection*. Oxford University Press on Demand.

Gettelman, A., J. Kay, and K. Shell, 2012: The evolution of climate sensitivity and climate feedbacks in the community atmosphere model. *Journal of Climate*, **25 (5)**, 1453–1469.

Golaz, J.-C., V. E. Larson, and W. R. Cotton, 2002a: A PDF-based model for boundary layer clouds. part I: Method and model description. *Journal of the atmospheric sciences*, **59 (24)**, 3540–3551.

Golaz, J.-C., V. E. Larson, and W. R. Cotton, 2002b: A PDF-based model for boundary layer clouds. part II: Model results. *Journal of the atmospheric sciences*, **59 (24)**, 3552–3571.

Goren, T. and D. Rosenfeld, 2014: Decomposing aerosol cloud radiative effects into cloud cover, liquid water path and twomey components in marine stratocumulus. *Atmospheric Research*, **138**, 378–393.

Griffin, B. M. and V. E. Larson, 2013: Analytic upscaling of a local microphysics scheme. part II: Simulations. *Quarterly Journal of the Royal Meteorological Society*, **139 (670)**, 58–69.

Gryspeerdt, E., P. Stier, and D. Partridge, 2014: Satellite observations of cloud regime development: the role of aerosol processes. *Atmospheric Chemistry and Physics*, **14 (3)**, 1141–1158.

Guo, H., J.-C. Golaz, L. Donner, B. Wyman, M. Zhao, and P. Ginoux, 2015: CLUBB as a unified cloud parameterization: opportunities and challenges. *Geophysical Research Letters*, **42 (11)**, 4540–4547.

Guo, H., J.-C. Golaz, L. J. Donner, P. Ginoux, and R. S. Hemler, 2014: Multi-variate probability density functions with dynamics in the gfdl atmospheric general circulation model: Global tests. *Journal of Climate*, **27 (5)**, 2087–2108.

Hartmann, D., M. Ockert-Bell, and M. Michelsen, 1992: The effect of cloud type on Earth's energy balance: Global analysis. *J. Climate*, **5**, 1281–1304.

Heinze, R., et al., 2016: Large-eddy simulations over germany using ICON: A comprehensive evaluation. *to be submitted to Q. J. Roy. Meteor. Soc.*

Heise, E., B. Ritter, and E. Schrodin, 2006: Operational implementation of the multi-layer soil model TERRA. Tech. rep., Deutscher Wetterdienst: Offenbach, Germany. http://www.cosmo-model.org (accessed 9 May 2016).

Horváth, Á., C. Seethala, and H. Deneke, 2014: View angle dependence of MODIS liquid water path retrievals in warm oceanic clouds. *Journal of Geophysical Research: Atmospheres*, **119 (13)**, 8304–8328.

Jakob, C., 2003: An improved strategy for the evaluation of cloud parameterizations in GCMs. *Bulletin of the American Meteorological Society*, **84 (10)**, 1387.

Jakob, C., 2014: Going back to basics. *Nature Climate Change*, **4 (12)**, 1042–1045.

Kawai, H. and J. Teixeira, 2012: Probability density functions of liquid water path and total water content of marine boundary layer clouds: Implications for cloud parameterization. *Journal of Climate*, **25 (6)**, 2162–2177.

Khairoutdinov, M. and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Monthly weather review*, **128 (1)**, 229–243.

King, M., et al., 2003: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from modis. *IEEE Trans. Geosci. Rem. Sens.*, **41**, 442–458.

King, M. D., Y. J. Kaufman, W. P. Menzel, and D. Tanre, 1992: Remote sensing of cloud, aerosol, and water vapor properties from the moderate resolution imaging spectrometer (MODIS). *Geoscience and Remote Sensing, IEEE Transactions on*, **30 (1)**, 2–27.

Klein, S. and D. Hartmann, 1993: The seasonal cycle of low stratiform clouds. *J. Climate*, **6**, 1587–1606.

Köhler, M., 2005: Improved prediction of boundary layer clouds. *ECMWF Newsletter*, **104**, 18–22.

Koren, I., L. Oreopoulos, G. Feingold, L. Remer, and O. Altaratz, 2008: How small is a small cloud? *Atmospheric Chemistry and Physics*, **8 (14)**, 3855–3864.

Kubar, T., D. Hartmann, and R. Wood, 2007: Radiative and convective driving of tropical high clouds. *Journal of Climate*, **20 (22)**, 5510–5526.

Larson, V. and D. Schanen, 2013: The subgrid importance latin hypercube sampler (SILHS): A multivariate subcolumn generator. *Geoscientific Model Development*, **6 (5)**, 1813–1829.

Larson, V. E. and J.-C. Golaz, 2005: Using probability density functions to derive consistent closure relationships among higher-order moments. *Monthly weather review*, **133 (4)**.

Larson, V. E., J.-C. Golaz, and W. R. Cotton, 2002: Small-scale and mesoscale variability in cloudy boundary layers: Joint probability density functions. *Journal of the Atmospheric Sciences*, **59 (24)**, 3519–3539.

Larson, V. E. and B. M. Griffin, 2013: Analytic upscaling of a local microphysics scheme. part I: Derivation. *Quarterly Journal of the Royal Meteorological Society*, **139 (670)**, 46–57.

Larson, V. E., D. P. Schanen, M. Wang, M. Ovchinnikov, and S. Ghan, 2012: PDF parameterization of boundary layer clouds in models with horizontal grid spacings from 2 to 16 km. *Monthly Weather Review*, **140 (1)**.

Larson, V. E., R. Wood, P. R. Field, J.-C. Golaz, T. H. V. Haar, and W. R. Cotton, 2001a: Small-scale and mesoscale variability of scalars in cloudy boundary layers: One-dimensional probability density functions. *Journal of the atmospheric sciences*, **58 (14)**, 1978–1994.

Larson, V. E., R. Wood, P. R. Field, J.-C. Golaz, T. H. Vonder Haar, and W. R. Cotton, 2001b: Systematic biases in the microphysics and thermodynamics of numerical models that ignore subgrid-scale variability. *Journal of the atmospheric sciences*, **58 (9)**, 1117–1128.

Le Treut, H. and Z.-X. Li, 1991: Sensitivity of an atmospheric general circulation model to prescribed SST changes: Feedback effects associated with the simulation of cloud optical properties. *Climate Dynamics*, **5 (3)**, 175–187.

L'Ecuyer, P. and S. Côté, 1991: Implementing a random number package with splitting facilities. *ACM Transactions on Mathematical Software (TOMS)*, **17 (1)**, 98–111.

Lewellen, W. and S. Yoh, 1993: Binormal model of ensemble partial cloudiness. *Journal of the atmospheric sciences*, **50 (9)**, 1228–1237.

Lock, A., A. Brown, M. Bush, G. Martin, and R. Smith, 2000: A new boundary layer mixing scheme. part I: Scheme description and single-column model tests. *Monthly weather review*, **128 (9)**, 3187–3199.

Lott, F. and M. J. Miller, 1997: A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, **123 (537)**, 101–127.

Majewski, D., et al., 2002: The operational global icosahedral-hexagonal gridpoint model GME: Description and high-resolution tests. *Monthly Weather Review*, **130 (2)**, 319–338.

Malkus, J. S., 1958: On the structure of the trade wind moist layer. *Massachusetts Institute of Technology and Woods Hole Oceanographic Institution.*

Medeiros, B., B. Stevens, I. M. Held, M. Zhao, D. L. Williamson, J. G. Olson, and C. S. Bretherton, 2008: Aquaplanets, climate sensitivity, and low clouds. *Journal of Climate*, **21 (19)**, 4974–4991.

Mellor, G. L., 1977: The gaussian cloud model relations. *Journal of the Atmospheric Sciences*, **34 (2)**, 356–358.

Menzel, W., R. Frey, and B. Baum, 2015: Aqua MODIS cloud product 5-min L2 swath 1 km and 5 km, C6, NASA level-1 atmosphere archive and distribution system (LAADS) distributed active archive center (DAAC). Tech. rep., Goddard Space Flight Center, Greenbelt. doi:http://dx.doi.org/10.5067/MODIS/MOD06_L2.006.

Menzel, W. P., R. Frey, and B. Baum, 2016: *Cloud top properties and cloud phase algorithm theoretical basis document - version 11.* University of Wisconsin - Madison.

Merk, D., H. Deneke, B. Pospichal, and P. Seifert, 2015: Investigation of the adiabatic assumption for estimating cloud micro-and macrophysical properties from satellite and ground. *Atmospheric Chemistry and Physics Discussions*, **15 (4)**, 5129–5173.

Meyer, K. and S. Platnick, 2010: Utilizing the MODIS 1.38 μm channel for cirrus cloud optical thickness retrievals: Algorithm and retrieval uncertainties. *Journal of Geophysical Research: Atmospheres*, **115 (D24)**.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: Rrtm, a validated correlated-k model for the longwave. *Journal of Geophysical Research: Atmospheres (1984–2012)*, **102 (D14)**, 16 663–16 682.

Muhlbauer, A., I. McCoy, and R. Wood, 2014: Climatology of stratocumulus cloud morphologies: microphysical properties and radiative effects. *Atmospheric Chemistry and Physics Discussions*, **14 (5)**, 6981–7023.

Muller, C. and S. Bony, 2015: What favors convective aggregation and why? *Geophysical Research Letters*, **42 (13)**, 5626–5634.

Nakajima, T. and M. D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. part I: Theory. *Journal of the Atmospheric Sciences*, **47 (15)**, 1878–1893.

Naumann, A., A. Seifert, and J. Mellado, 2013: A refined statistical cloud closure using double-gaussian probability density functions. *Geoscientific Model Development*, **6 (5)**, 1641–1657.

NCL, 2016: NCAR command language (version 6.3.0) [software]. Tech. rep., Boulder, Colorado: UCAR/NCAR/CISL/TDD. http://dx.doi.org/10.5065/D6WD3XH5.

Neggers, R., 2015a: Attributing the behavior of low-level clouds in large-scale models to subgrid-scale parameterizations. *Journal of Advances in Modeling Earth Systems*.

Neggers, R., 2015b: Exploring bin-macrophysics models for moist convective transport and clouds. *Journal of Advances in Modeling Earth Systems.*

Neggers, R., H. Jonker, and A. Siebesma, 2003: Size statistics of cumulus cloud populations in large-eddy simulations. *Journal of the atmospheric sciences*, **60 (8)**, 1060–1074.

Neggers, R., A. Siebesma, and T. Heus, 2012: Continuous single-column model evaluation at a permanent meteorological supersite. *Bulletin of the American Meteorological Society*, **93 (9)**, 1389–1400.

Neggers, R., A. Siebesma, G. Lenderink, and A. Holtslag, 2004: An evaluation of mass flux closures for diurnal cycles of shallow cumulus. *Monthly weather review*, **132 (11)**, 2525–2538.

Neggers, R. A., M. Köhler, and A. C. Beljaars, 2009: A dual mass flux framework for boundary layer convection. part I: Transport. *Journal of the Atmospheric Sciences*, **66 (6)**, 1465–1487.

Neggers, R. A. J., 2009: A dual mass flux framework for boundary layer convection. part II: Clouds. *Journal of the Atmospheric Sciences*, **66**, 1489–1506.

Nuijens, L., B. Medeiros, I. Sandu, and M. Ahlgrimm, 2015a: The behavior of trade-wind cloudiness in observations and models: The major cloud components and their variability. *Journal of Advances in Modeling Earth Systems.*

Nuijens, L., B. Medeiros, I. Sandu, and M. Ahlgrimm, 2015b: Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade wind layer. *Journal of Advances in Modeling Earth Systems.*

Oreopoulos, L. and R. F. Cahalan, 2005: Cloud inhomogeneity from MODIS. *Journal of climate*, **18 (23)**, 5110–5124.

Oreopoulos, L. and R. Davies, 1998: Plane parallel albedo biases from satellite observations. part I: Dependence on resolution and other factors. *Journal of climate*, **11 (5)**, 919–932.

Orr, A., P. Bechtold, J. Scinocca, M. Ern, and M. Janiskova, 2010: Improved middle atmosphere climate and forecasts in the ECMWF model through a nonorographic gravity wave drag parameterization. *Journal of Climate*, **23 (22)**, 5905–5926.

Perraud, E., F. Couvreux, S. Malardel, C. Lac, V. Masson, and O. Thouron, 2011: Evaluation of statistical distributions for the parametrization of subgrid boundary-layer clouds. *Boundary-layer meteorology*, **140 (2)**, 263–294.

Pincus, R., R. Hemler, and S. A. Klein, 2006: Using stochastically generated subcolumns to represent cloud structure in a large-scale model. *Monthly weather review*, **134 (12)**, 3644–3656.

Pincus, R. and S. A. Klein, 2000: Unresolved spatial variability and microphysical process rates in large-scale models. *Journal of Geophysical Research: Atmospheres*, **105 (D22)**, 27 059–27 065.

Pincus, R., S. A. McFarlane, and S. A. Klein, 1999: Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites. *Journal of geophysical research*, **104 (D6)**, 6183–6191.

Platnick, S., S. Ackerman, M. King, and et al., 2015: Atmosphere level 2 cloud product (06 l2). NASA MODIS adaptive processing system. *Goddard Space Flight Center, USA*, URL `http://dx.doi.org/10.5067/MODIS/MOD06_L2.006`.

Platnick, S., M. King, S. Ackerman, W. Menzel, B. Baum, J. Riédi, and R. Frey, 2003: The modis cloud products: Algorithms and examples from terra. *IEEE Trans. Geosci. Rem. Sens.*, **41**, 459–473.

Price, J. D. and R. Wood, 2002: Comparison of probability density functions for total specific humidity and saturation deficit humidity, and consequences for cloud parametrization. *Quarterly Journal of the Royal Meteorological Society*, **128 (584)**, 2059–2072.

Prill, F., 2014: DWD icon tools documentation. Tech. rep., Deutscher Wetterdienst (DWD).

Quaas, J., 2012: Evaluating the critical relative humidity as a measure of subgrid-scale variability of humidity in general circulation model cloud cover parameterizations using satellite data. *Journal of Geophysical Research*, **117 (D9)**, D09 208.

Räisänen, P., H. W. Barker, M. F. Khairoutdinov, J. Li, and D. A. Randall, 2004: Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quarterly Journal of the Royal Meteorological Society*, **130 (601)**, 2047–2067.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski, 2003: Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society*, **84 (11)**, 1547–1564.

Rieck, M., C. Hohenegger, and C. C. van Heerwaarden, 2014: The influence of land surface heterogeneities on cloud size development. *Monthly Weather Review*, **142 (10)**, 3830–3846.

Rosch, J., T. Heus, M. Brueck, M. Salzmann, J. Mülmenstädt, L. Schlemmer, and J. Quaas, 2015: Analysis of diagnostic climate model cloud parameterisations using large-eddy simulations. *Quarterly Journal of the Royal Meteorological Society*, doi: 10.1002/qj.2515.

Saha, S., et al., 2010: The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, **91 (8)**, 1015–1057.

Sakradzija, M., V. Lucarini, and A. Seifert, 2015: A stochastic parameterization of shallow cumulus convection for high-resolution numerical weather and climate models. Ph.D. thesis, Universität Hamburg Hamburg.

Schemann, V., B. Stevens, V. Grützun, and J. Quaas, 2013: Scale dependency of total water variance and its implication for cloud parameterizations. *Journal of the Atmospheric Sciences*, **70 (11)**, 3615–3630.

Seethala, C. and Á. Horváth, 2010: Global assessment of AMSR-E and MODIS cloud liquid water path retrievals in warm oceanic clouds. *Journal of Geophysical Research: Atmospheres*, **115 (D13)**.

Seifert, A. and T. Heus, 2013: Large-eddy simulation of organized precipitating trade wind cumulus clouds. *Atmos. Chem. Phys*, **13**, 5631–5645.

Sherwood, S. C., S. Bony, and J.-L. Dufresne, 2014: Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, **505 (7481)**, 37–42.

Siebesma, A. P., P. M. Soares, and J. Teixeira, 2007: A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *Journal of the atmospheric sciences*, **64 (4)**.

Siebesma, A. P., et al., 2004: Cloud representation in general-circulation models over the northern pacific ocean: A eurocs intercomparison study. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3245–3267.

Smith, R., 1990: A scheme for predicting layer clouds and their water content in a general circulation model. *Quarterly Journal of the Royal Meteorological Society*, **116 (492)**, 435–460.

Soares, P., P. Miranda, A. Siebesma, and J. Teixeira, 2004: An eddy-diffusivity/mass-flux parametrization for dry and shallow cumulus convection. *Quarterly Journal of the Royal Meteorological Society*, **130 (604)**, 3365–3383.

Sommeria, G. and J. Deardorff, 1977: Subgrid-scale condensation in models of nonprecipitating clouds. *Journal of the Atmospheric Sciences*, **34 (2)**, 344–355.

Staniforth, A. and J. Thuburn, 2012: Horizontal grids for global weather and climate prediction models: a review. *Quarterly Journal of the Royal Meteorological Society*, **138 (662)**, 1–26.

Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *Journal of climate*, **18 (2)**, 237–273.

Stevens, B., 2005: Atmospheric moist convection. *Annu. Rev. Earth Planet Sci.*, **33**, 605–643, doi:10.1146/annurev.earth.33.092203.122658.

Stevens, B. and S. Bony, 2013: Waterin the atmosphere. *Phys. Today*, **66 (6)**, 29.

Stevens, B., et al., 2005: Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly weather review*, **133 (6)**, 1443–1462.

Stevens, B., et al., 2013: Atmospheric component of the MPI-M earth system model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, **5 (2)**, 146–172.

Sundqvist, H., E. Berge, and J. Kristjánsson, 1989: Condensation and cloud parameterization studies with a mesoscale numerical weather prediction model. *Monthly Weather Review*, **117 (8)**, 1641–1657.

Teixeira, J., et al., 2011: Tropical and subtropical cloud transitions in weather and climate prediction models: The GCSS/WGNE pacific cross-section intercomparison (GPCI). *Journal of Climate*, **24 (20)**, 5223–5256.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, **117 (8)**, 1779–1800.

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Monthly Weather Review*, **121 (11)**, 3040–3061.

Tobin, I., S. Bony, and R. Roca, 2012: Observational evidence for relationships between the degree of aggregation of deep convection, water vapor, surface fluxes, and radiation. *Journal of Climate*, **25 (20)**, 6885–6904.

Tomassini, L., A. Voigt, and B. Stevens, 2015: On the connection between tropical circulation, convective mixing, and climate sensitivity. *Quarterly Journal of the Royal Meteorological Society*, **141 (689)**, 1404–1416.

Tompkins, A., 2002: A prognostic parameterization for the subgrid-scale variability of water vapor and clouds in large-scale models and its use to diagnose cloud cover. *Journal of the Atmospheric Sciences*, **59 (12)**, 1917–1942.

Tompkins, A., 2003: Impact of temperature and humidity variability on cloud cover assessed using aircraft data. *Quarterly Journal of the Royal Meteorological Society*, **129 (592)**, 2151–2170.

Tompkins, A. M., 2008: Cloud parametrization. Tech. rep., ECMWF Seminar on Parametrization of Subgrid Physics Processes.

van der Dussen, J., S. de Roode, and A. Siebesma, 2016: How large-scale subsidence affects stratocumulus transitions. *Atmospheric Chemistry and Physics*, **16 (2)**, 691–701.

Vial, J., J. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Clim. Dynam.*, 1–24, doi:DOI10.1007/s00382-013-1725-9.

Wan, H., et al., 2013: The ICON-1.2 hydrostatic atmospheric dynamical core on triangular grids, part I: formulation and performance of the baseline version. *Geoscientific Model Development*, **6**, 735–763.

Warren, S., C. Hahn, J. London, R. Chervin, and R. Jenne, 1988: Global distribution of total cloud cover and cloud type amounts over the ocean. Tech. rep., USDOE Office of Energy Research, Washington, DC (USA). Carbon Dioxide Research Div.; National Center for Atmospheric Research, Boulder, CO (USA). doi:10.2172/5415329.

Webb, M., et al., 2006: On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, **27 (1)**, 17–38.

Weber, T., J. Quaas, and P. Räisänen, 2011: Evaluation of the statistical cloud scheme in the ECHAM5 model using satellite data. *Quarterly Journal of the Royal Meteorological Society*, **137 (661)**, 2079–2091.

Wood, R., 2012: Stratocumulus clouds. *Monthly Weather Review*, **140 (8)**, 2373–2423.

Wood, R. and C. Bretherton, 2006: On the relationship between stratiform low cloud cover and lower-tropospheric stability. *J. Climate*, **19**, 6425–6432.

Wood, R. and P. R. Field, 2011: The distribution of cloud horizontal sizes. *Journal of Climate*, **24 (18)**, 4800–4816.

Wood, R. and D. L. Hartmann, 2006: Spatial variability of liquid water path in marine low cloud: The importance of mesoscale cellular convection. *Journal of climate*, **19 (9)**, 1748–1764.

Zängl, G., 2013: ICON: The icosahedral nonhydrostatic modelling framework of DWD and MPI-M. *Proc. ECMWF Seminar on Numerical Methods for Atmosphere and Ocean Modelling.*

Zängl, G., D. Reinert, P. Rípodas, and M. Baldauf, 2015: The ICON (icosahedral non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, **141 (687)**, 563–579.

Zelinka, M. D., S. A. Klein, and D. L. Hartmann, 2012: Computing and partitioning cloud feedbacks using cloud property histograms. part I: Cloud radiative kernels. *Journal of Climate*, **25 (11)**, 3715–3735.

Zhao, G. and L. Di Girolamo, 2006: Cloud fraction errors for trade wind cumuli from eos-terra instruments. *Geophys. Res. Lett.*, **33 (L20802)**, doi:DOI:10.1029/2006GL027088.

Zhao, M., I. M. Held, S.-J. Lin, and G. A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50-km resolution GCM. *Journal of Climate*, **22 (24)**, 6653–6678.

List of Figures

List of Tables

ACKNOWLEDGEMENTS

First, I want to thank Johannes Quaas for being a great advisor and for giving me the opportunity to develop this project on basis of his initial ideas. He continuously inspired and guided my efforts with his vast big picture view, which supported me to become an independent researcher. Special thanks to Chris Golaz who enabled my research visit to the GFDL and warmly welcomed me. I am very thankful for his great scientific input and encouraging spirit.

I deeply want to thank my whole working group ‘Clouds and Global Climate’ at the Institute for Meteorology in the Univerität Leipzig, especially Edward Gryspeerdt, Karoline Block, Tom Goren, Johannes Mülmenstädt, Christine Nam, Marc Salzmann, Birgit Seydel, Odran Sourdeval, Dipu Sudhakar, Claudia Unglaub, Aswathy Nair and Falk Kaiser for all the productive discussions, help and encouragement over the last three years. You all made my time in Leipzig pleasant and scientifically interesting.

In the same I manner thank Jörg Burdanowitz, Louise Nuijens, Bjorn Stevens and Angela Gruber, Ann-Kirsitn Naumann and Mirjana Sakradzija for making my time at the Max Planck Institute for Meteorology so pleasant.

Furthermore I would like to express my thanks to Huan Guo, Leo Donner, Yi Ming and Lisa Allan you all inspired, helped and focused me during my time in Princeton and the Geophysical Fluid Dynamics Laboratory (GFDL). I am also grateful to Hartwig Deneke and Sebastian Bley for valuable comments on satellite data processing. Thanks also to Beate Richter and Andreas Macke for their organizational effort and support for the graduate school on ‘Clouds, Aerosols and Radiation: Mineral Dust’. Thanks to Martin Köhler and Daniel Klocke for many helpful comments and providing ICON input data. For widening my view on my topic I would also like to thank Roel Neggers, Axel Seifert and Thijs Heus you gave me further inspiration. Furthermore I am thankful for the amazing computational infrastructure of the German Climate Computing Centre (DKRZ) and the Oak Ridge National Laboratory for their effort which provided the basis for the GCM simulations. The Aqua/MODIS data was acquired from the Level-1 Atmosphere Archive and Distribution System (LAADS) Distributed Active Archive Center (DAAC) of the NASA Goddard Space Flight Center. This work is financially supported by the Federal Ministry of Education and Research in Germany (BMBF) through the research program High Definition Clouds and Precipitation for Climate Prediction - HD(CP)^2 (FKZ: 01LK1210D, 01LK1213B and 01LK1218A), and by the European Research Council trough the research program QUAERERE (grant agreement no. 306284). Throughout this work the HDCP2 community created a very stimulating environment which was foremost facilitated by Florian Rauser. Also I would like to thank the DAAD for their financial support during this period. I thank everyone, who directly or indirectly influenced my research during the last three years! In particular, I would like to thank my family, friends and specifically Nina for their enthusiastic support during my entire studies.

LIST OF PUBLICATIONS

Wissenschaftliche Veröffentlichungen in Zusammenhang mit dieser Arbeit

Scientific publications in context to this work

Brueck, M., L. Nuijens, and B. Stevens, 2015: On the seasonal and synoptic time scale variability of the north Atlantic trades and its low-level clouds. J. Atmos. Sci., 72 (4), doi:JAS-D-14-0054.1

Rosch, J., T. Heus, **M. Brueck**, M. Salzmann, J. Mülmenstädt, L. Schlemmer, and J. Quaas, 2015: Analysis of diagnostic climate model cloud parameterisations using large-eddy simulations. Q. J. Roy. Meteor. Soc., doi:10.1002/qj.2515

Heinze, R. A. Dipankar, C. Carbajal-Henken, C. Moseley, O. Sourdeval, S. Trömel, X. Xie, P. Adamidis, F. Ament, C. Barthlott, A. Behrendt, U. Blahak, S. Bley, S. Brdar, **M. Brueck**, S. Crewell, H. Deneke, P. Di Girolamo, R. Evaristo, C. Frank, K. Gorges, M. Hanke, A. Hansen, C. Hoose, T. Jahns, N. Kalthoff, D. Klocke, S. Kneifel, P. Knippertz, A. Kuhn, T. van Laar, A. Macke, V. Maurer, C. Meyer, S. Muppa, R. Neggers, F. Pantillon, B. Pospichal, N. Röber, L. Scheck, A. Seifert, F. Senf, P. Siligam, C. Simmer, S. Steinke, B. Stevens, K. Wapler, M. Weniger, V. Wulfmeyer, G. Zängl, D. Zhang and J. Quaas (2016): Large-eddy simulations over Germany using ICON: A comprehensive evaluation. Q. J. Roy. Meteor. Soc. (accepted)

Brueck, M., J. Quaas, J.C. Golaz and H. Guo, 2016: Importance of subgrid-scale temperature variability for cloud parameterizations. Mon. Wea. Rev. (submitted)

Brueck, M., J. Quaas, M. Köhler and D. Klocke, 2016: Process-based evaluation of cloud subgrid-scale variability in the ICON-GCM model. J. Adv. Model. Earth. Syst., in preparation